Chapter 21

# Connectedness

# Introduction

Welcome to the final book of M303. In this book we continue our investigation of metric spaces, but first we pause to take stock. So far you have studied number theory, groups, rings, fields and metric spaces. It is natural to see these as very different topics, but you may be starting to get a sense of an overall similarity. When studying groups, rings and fields, we start by defining a set, together with an operation or two upon that set. We then require that certain rules are satisfied. We often add additional conditions (such as requiring that a group operation is commutative). The integers form a ring, and so we can view number theory as studying one specific ring. A metric space may seem like a very different animal, but the process is the same: we start with a set of elements (such as points on the real line) and add some structure (an idea of 'distance').

$\mathbb{Z}$ is a particularly worthwhile ring to study as there is always a unique ring homomorphism from $\mathbb{Z}$ into any ring with an identity.

In general, algebraists talk of a 'set of elements' and geometers speak of a 'space of points'; but the distinction is somewhat artificial. You may have studied linear algebra and noticed the algebraic concept of addition appearing alongside the geometric notion of a modulus.

There are many mathematical objects which combine algebraic and geometric ideas, such as Lie groups, in which the group operation is required to be, in some sense, continuous.

We can apply varying structures to the same set or space. For example, we can consider $\mathbb{Z}$, the set of integers, as an abelian group, a ring or as a metric space (with an appropriate metric such as the Euclidean metric or the $p$-adic metric).

When studying mathematical objects, we need to work out what we mean by a different object. For example, we study groups up to isomorphism. We start this chapter by discussing what we mean by 'different' in the context of a metric space.

Our goal in this chapter is to understand what it means to say that a metric space is connected. In doing so, we discuss two different notions of connectedness and discover how they are preserved (or not) under particular transformations of metric spaces known as homeomorphisms.

As a result of these notions of connectedness, we are able to fulfil the first of our promises from Book D, Chapter 13 and prove a version of the Intermediate Value Theorem for continuous functions from a metric space $(X, d)$ to $\mathbb{R}$.

## Chapter guide

The key ideas here are those of homeomorphisms, connectedness and path-connectedness. We look at homeomorphisms in Section 1, connectedness in Section 3 and path-connectedness in Section 5. Section 4 looks at connectedness in Euclidean space in particular, and contains a generalisation of the Intermediate Value Theorem. The final section contains an example which illustrates that a connected space need not be path-connected. You should try to understand the ideas underlying this example but you do not need to learn the technical details of the construction.

# 1 Homeomorphisms

Two groups are isomorphic if we can find a bijection between the two sets which preserves the effect of the group operation. We want a similar idea to describe when two metric spaces $(X, d_X)$ and $(Y, d_Y)$ are 'essentially the same'. Our choice is motivated by a desire to study continuity, in its purest form, and so we use the existence of a bijective $(d_X, d_Y)$-continuous map, with a $(d_Y, d_X)$-continuous inverse, between the two spaces to define our idea of equivalence.

**Definition 1.1** *Homeomorphism*

Let $(X, d)$ and $(Y, e)$ be two metric spaces. Then $(X, d)$ and $(Y, e)$ are **homeomorphic** metric spaces if there exists a bijection $f\colon X \to Y$ such that $f\colon X \to Y$ is $(d, e)$-continuous and $f^{-1}\colon Y \to X$ is $(e, d)$-continuous. We say that $f$ is a **homeomorphism**.

Topology (a generalisation of metric spaces mentioned in Book D, Chapter 16) is often called 'rubber sheet geometry' for this reason.

Intuitively, two spaces are homeomorphic if one can be continually deformed into the other. We can twist and stretch our spaces, but we cannot create holes (which would create a discontinuity).

The Composition and Restriction Rules for continuous functions generalise as expected to homeomorphisms. In particular, we have the following result.

**Proposition 1.2** *Restriction Rule for homeomorphisms*

Let $h\colon X \to Y$ be a homeomorphism between two metric spaces $(X, d)$ and $(Y, e)$, and let $X' \subseteq X$. Then $h|_{X'}$, the restriction of $h$ to $X'$, gives a homeomorphism from $X'$ to the image $h(X') \subseteq Y$ so that $(X', d_{X'})$ and $(h(X'), e_{h(X')})$ are homeomorphic metric spaces.

The induced metric, $d_A$, for a subset $A$ of a metric space $(X, d)$ is given by $d_A(a, b) = d(a, b)$ for $a, b \in A$.

**Proof** We need to show that $h|_{X'}$ is $(d_{X'}, e_{h(X')})$-continuous, is a bijection, and has a $(e_{h(X')}, d_{X'})$-continuous inverse. As $h$ is $(d, e)$-continuous and $h^{-1}$ is $(e, d)$-continuous, the Restriction Rule for continuous functions (Book D, Chapter 15, Theorem 1.6) shows that $h|_{X'}$ is $(d_{X'}, e_{h(X')})$-continuous and $h^{-1}|_{h(X')}$ is $(e_{h(X')}, d_{X'})$-continuous. Clearly $h|_{X'}$ and $h^{-1}|_{h(X')}$ are one–one (since $h$ and $h^{-1}$ are) and, by definition, $h|_{X'}$ is onto its image and $h^{-1}|_{h(X')}$ is onto its preimage. ■

---

**Example 1.3**

Consider the sets $X = (0, 1)$ and $Y = (0, 5)$ (each with the metric induced by the Euclidean metric on $\mathbb{R}$). Then the map $f\colon x \mapsto 5x$ gives a homeomorphism from $X$ to $Y$, and so $(X, d_X^{(1)})$ and $(Y, d_Y^{(1)})$ are homeomorphic metric spaces.

### Example 1.4

Consider the sets $X = [1, 2]$ and $Y = [-10, 10]$ (each with the metric induced by the Euclidean metric on $\mathbb{R}$). Then the map $g\colon X \to Y$ given by $g\colon x \mapsto -10 + 20(x - 1)$ is a homeomorphism from $X$ to $Y$, and so $(X, d_X^{(1)})$ and $(Y, d_Y^{(1)})$ are homeomorphic metric spaces.

### Example 1.5

Let $S$ be the square $\{(x, y) \in \mathbb{R}^2 \colon 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1\}$ and $R$ be the rectangle $\{(x, y) \in \mathbb{R}^2 \colon 0 \leq x \leq 2 \text{ and } 0 \leq y \leq 1\}$, each with the metric induced by the Euclidean metric on $\mathbb{R}^2$. The map $h\colon S \to R$ given by $h\colon (x, y) \mapsto (2x, y)$ gives a homeomorphism from $S$ to $R$, and so $(S, d_S^{(2)})$ and $(R, d_R^{(2)})$ are homeomorphic metric spaces.

### Exercise 1.1

Let $A = [0, 1]$ and $B = [0, n]$ for some $n \in \mathbb{N}$, each with the metric induced by the Euclidean metric on $\mathbb{R}$. Show that $(A, d_A^{(1)})$ and $(B, d_B^{(1)})$ are homeomorphic metric spaces by writing down a specific homeomorphism from $A$ to $B$. Can $n$ be any real number?

### Exercise 1.2

Let $X = [a, b]$ and $Y = [c, d]$, each with the metric induced by the Euclidean metric on $\mathbb{R}$. By writing down an explicit homeomorphism between $X$ and $Y$, show that any two closed intervals $[a, b]$ and $[c, d]$ with $a \neq b$ and $c \neq d$, and each with the Euclidean metric, are homeomorphic. Does it matter if $a < b$ and $c > d$?

### Exercise 1.3

Let $X = [a, b)$ and $Y = [c, d)$, each with the metric induced by the Euclidean metric on $\mathbb{R}$. Show that if $a \neq b$ and $c \neq d$ then $X$ and $Y$ are homeomorphic metric spaces. Is the same true for $(a, b]$ and $[c, d)$?

Instead of writing down homeomorphisms explicitly (especially when working in Euclidean space), we often refer to the 'obvious' map. However, it is important to be sure that 'the obvious map' is a continuous bijection with a continuous inverse!

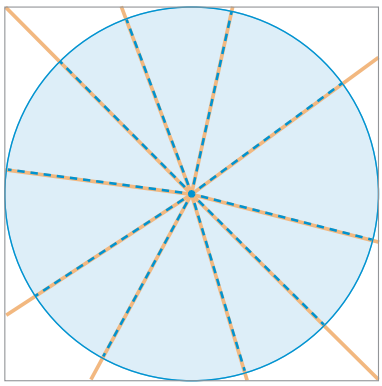


**Figure 1.1**

### Example 1.6

The closed disc, $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, and the square $S = \{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 1 \text{ and } -1 \leq y \leq 1\}$, each with the metric induced by the Euclidean metric on $\mathbb{R}^2$, are homeomorphic metric spaces. The homeomorphism is given by mapping the radial lines of the disc onto the corresponding radial lines of the square; the function 'stretches' the radial lines of the disc so that they map in a one–one way onto the radial lines of the square (see Figure 1.1).

### Example 1.7

The unit ball in $\mathbb{R}^3$, $B_{d(3)}[(0, 0, 0), 1]$ and the cube $\{(x, y, z) \in \mathbb{R}^3 : -1 \leq x \leq 1 \text{ and } -1 \leq y \leq 1 \text{ and } -1 \leq z \leq 1\}$, each with the metric induced by the Euclidean metric on $\mathbb{R}^3$, are homeomorphic metric spaces. The homeomorphism is given by a three-dimensional version of the mapping shown in Figure 1.1.

### Exercise 1.4

Convince yourself that a coffee cup is homeomorphic to a doughnut (when each are viewed as subsets of $\mathbb{R}^3$ and given the metric induced by the Euclidean metric on $\mathbb{R}^3$).

Once we start to think about homeomorphisms, we sometimes find that the closed-set characterisation of continuity, given in Theorem 7.1 of Chapter 16, can be more useful than the sequential definition. We restate this characterisation here for reference.

### Theorem 1.8

Let $(X, d)$ and $(Y, e)$ be metric spaces and suppose that $f : X \to Y$. Then $f$ is $(d, e)$-**continuous** on $X$ if, and only if, for each $e$-closed set $E \subseteq Y$ $f^{-1}(E)$ is $d$-closed.

#### Remark

We could, of course, phrase this in terms of open sets rather than closed sets (using Book D, Chapter 16, Corollary 7.2).

We use this characterisation of continuity in the following exercise to demonstrate an unusual property of discrete metric spaces.

### Worked Exercise 1.9

Let $(X, d_0)$ be any space with the discrete metric. Let $(Y, d)$ be any metric space and $f\colon X \to Y$ be any function. Show that $f$ is $(d_0, d)$-continuous.

#### Solution

From Worked Exercise 1.11 of Chapter 16 we know that all subsets of a discrete metric space are $d_0$-closed. Therefore for any $d$-closed set $A \subseteq Y$, we must have that $f^{-1}(A)$ is $d_0$-closed in $X$, so $f$ is $(d_0, d)$-continuous by Theorem 7.1 of Chapter 16.

We have now proved the following.

### Proposition 1.10

Let $(X, d_0)$ be any space with the discrete metric. Let $(Y, d)$ be any metric space and $f\colon X \to Y$ be any function. Then $f$ is $(d_0, d)$-continuous.

Discrete metric spaces have rather unintuitive properties (and so are a good source of counter-examples to conjectures that might seem 'obviously true').

We defined equivalent metrics for the same underlying set in Definition 3.5 of Chapter 15. It is natural to wonder whether two spaces with equivalent metrics are homeomorphic. We ask you to determine this in the next worked exercise.

### Worked Exercise 1.11

Consider a set $X$ with two equivalent metrics $d$ and $d'$. Show that $(X, d)$ and $(X, d')$ are homeomorphic metric spaces.

#### Solution

We need to find a $(d, d')$-continuous bijective map, $f\colon X \to X$ with a $(d', d)$-continuous inverse. Consider the identity map $\mathrm{Id}\colon X \to X$. This is clearly bijective and $(d, d)$-continuous. Therefore by Theorem 3.8 of Chapter 15 it is $(d, d')$-continuous. The same result shows that the inverse, $\mathrm{Id}^{-1}\colon X \to X$, is $(d', d)$-continuous. Therefore $(X, d)$ and $(X, d')$ are homeomorphic metric spaces.

$\mathrm{Id}(x) = x$ for each $x \in X$.

Worked Exercise 1.11 shows that we cannot tell the difference between metric spaces with equivalent metrics when we are working up to homeomorphism. It turns out that many properties of metric spaces are preserved by homeomorphism. By this we mean that if $X$ and $Y$ are homeomorphic spaces, and $X$ has one of these properties, then $Y$ must also have the property. Properties preserved by homeomorphism are called *topological invariants*.

The collection of all the open sets in a metric space is called the topology on the space (Book D, Chapter 16, Definition 7.3).

**Definition 1.12** *Topological invariant*

We say that a property of metric spaces is a **topological invariant** if it is preserved by homeomorphisms. That is, if two spaces are homeomorphic and one has the property, then so must the other.

Connectedness, intuitively, tells us whether or not a space is 'in one piece'. Compactness is a concept which gives us an idea of 'size' for a metric space. Completeness, in some sense, tells us whether a space has any 'missing points'.

We will show that the two properties explored in the first two chapters of this book, namely connectedness and compactness, are both topological invariants but that the last of the 'three Cs', completeness, is not.

## 1.1 Homeomorphic subsets of $\mathbb{R}$: part 1

We finish this section by taking a first look at which intervals of $\mathbb{R}$ are homeomorphic.

First we note that if a set $A \subseteq \mathbb{R}$ is an interval, then whenever $a, b \in A$ and $c \in \mathbb{R}$ with $a < c < b$, we have $c \in A$. In particular, we can view $\mathbb{R}$ as an interval by thinking of it as $(-\infty, \infty)$. The empty set, $\emptyset = (a, a)$ and a single point $\{a\} = [a, a]$ are also intervals.

In this section we assume that all intervals have the metric induced by the Euclidean metric on $\mathbb{R}$, which we write as $d$. In Exercise 1.2 we used the map $f_I : x \mapsto c + \frac{d-c}{b-a}(x-a)$ to show that any two closed intervals of $\mathbb{R}$ are homeomorphic. The restriction of this map to the open interval $(a, b)$ shows that any two bounded open intervals are homeomorphic. We also used this map in Exercise 1.3 to show that any two bounded half-open intervals are homeomorphic. In other words, any two bounded intervals of the same 'type' are homeomorphic.

A *bounded* interval is one of the form $[a, b], [a, b), (a, b]$ or $(a, b)$ for $a, b \in \mathbb{R}$.

In the next exercise we ask you to show that the map $f_I$ does not give a homeomorphism between bounded intervals of different types.

**Exercise 1.5**

Show that $f : x \mapsto c + \frac{d-c}{b-a}(x-a)$ is not a homeomorphism from $(a, b]$ to $(c, d)$.

Of course the fact that $f_I$ does not give us an explicit homeomorphism between intervals of different types does not prove that no such map exists. Showing that two spaces are not homeomorphic can be difficult, and mathematicians often use properties that are preserved by homeomorphism to distinguish between such spaces. In Subsection 4.1 we use a topological invariant, *connectedness*, to prove that intervals of different types are not homeomorphic.

Sometimes we can show that two sets are not homeomorphic without resorting to topological invariants. We ask you to do this in the next exercise.

### Exercise 1.6

Show that the metric space $(\mathbb{R}, d)$, where $d$ is the Euclidean metric, and the set $\{1, 2\}$ with the discrete metric, are not homeomorphic metric spaces.

It may seem intuitively obvious that any two intervals of the same type are homeomorphic. What is more surprising is that the whole of $\mathbb{R}$ itself (with the Euclidean metric) and the interval $(0, 1)$ (with the metric induced by the Euclidean metric on $\mathbb{R}$) are homeomorphic metric spaces.

### Worked Exercise 1.13

Show that the metric space $(\mathbb{R}, d^{(1)})$ and the metric space $\big((0,1), d^{(1)}_{(0,1)}\big)$ are homeomorphic metric spaces.

#### Solution

To see this we consider the function $\phi\colon \mathbb{R} \to (0,1)$ given by $\phi(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x)$ (taking the principal value of arctan). This is the sum of basic continuous functions (so is $(d^{(1)}, d^{(1)}_{(0,1)})$-continuous) and has a $(d^{(1)}_{(0,1)}, d^{(1)})$-continuous inverse, namely $\tan\big(\pi(x - 1/2)\big)$. The fact that the inverse is defined on $(0,1)$ shows that $\phi$ is onto, and standard properties of arctan show that it is one–one.

### Remark

Don't worry if the map $\phi$ appeared in the above discussion like a rabbit out of a hat. Maps like this do not usually immediately spring to mind. They are constructed, often after considerable effort, by thinking about the kind of map that might work and then modifying it. In this case, we want a map that squashes all of the real line into $(0,1)$. We know that arctan squashes things up, so we use that and then fiddle around to get the scaling right.

### Exercise 1.7

Show that $(\mathbb{R}, d^{(1)})$ and $\left((a,\infty), d^{(1)}_{(a,\infty)}\right)$, where $a$ is any real number, are homeomorphic metric spaces.

To summarise, we know that any two bounded intervals of the same 'type' are homeomorphic, and that $\mathbb{R}$ is homeomorphic to the $d^{(1)}$-open interval $(0,1)$ (and so, since the composition of two homeomorphisms is also a homeomorphism, we know that $\mathbb{R}$ is homeomorphic to any non-empty $d^{(1)}$-open interval). The map $f\colon x \mapsto \tan(x)$ restricted to the interval $[0, \frac{\pi}{2})$ can be used to give a homeomorphism between $[0, \frac{\pi}{2})$ and $[0,\infty)$. That, combined with Exercise 1.3, and the fact that the composition of two homeomorphisms is a homeomorphism, shows that any two half-open intervals are homeomorphic.

Therefore we have now identified five classes of intervals:

- the empty set $\emptyset$
- the singleton sets $\{a\} = [a,a]$
- the open intervals $(a,b)$ for $a < b$, $(a,\infty)$ and $(-\infty,a)$ for $a \in \mathbb{R}$, and $\mathbb{R}$
- the closed intervals $[a,b]$ for $a < b$
- the intervals of the form $[a,b)$ and $(a,b]$ for $a < b$ and $[a,\infty)$ and $(-\infty,a]$.

Intervals in the same class are homeomorphic. We do not yet know whether intervals in different classes are not homeomorphic. We shall investigate this question in Subsection 4.1.

## 2 Closed and open sets revisited

Recall from Chapter 16 that a set, $E$, is $d$-closed for a metric $d$ if every $d$-convergent sequence $(x_n)$, with terms $x_n \in E$, has its limit, $x$, in $E$; and that a $d$-open set is the complement of a $d$-closed set.

When examining a *subspace* $E$ of a metric space $(X,d)$ we must be a little careful. Consider the interval $(0,1)$. Stop for a moment (not for too long!) and think about whether this interval is closed ....

You may have decided that this interval is not closed – we tend to refer to intervals like this as 'open intervals'. The interval $(0,1)$ is certainly not $d^{(1)}$-closed as a subset of $\mathbb{R}$ with the Euclidean metric $d^{(1)}$, but is it $d^{(1)}_{(0,1)}$-closed as a subset of $(0,1)$ with the induced metric $d^{(1)}_{(0,1)}$? In other words, what happens if we consider the set $(0,1)$, with the subspace metric, $d^{(1)}_{(0,1)}$, as our whole space? We know from Chapter 16, Lemma 1.2 that the whole space, $X$, is always a $d$-closed subset of a metric space $(X,d)$, so if we consider the set $(0,1)$ as a subset of the metric space

$\left((0,1), d^{(1)}_{(0,1)}\right)$ then it is $d^{(1)}_{(0,1)}$-closed. So, whether or not a particular set is closed depends on 'where this set is' (and of course on which metric we are using). The interval $(0,1)$ is not $d^{(1)}$-closed as a subset of $\left(\mathbb{R}, d^{(1)}\right)$, but is $d^{(1)}_{(0,1)}$-closed as a subset of $\left((0,1), d^{(1)}_{(0,1)}\right)$.

It is natural to ask whether there is a space, $(X,d)$, containing $(0,1)$ as a *proper* subset in which $(0,1)$ is $d$-closed. In order to answer this, it is helpful to use the characterisation of closed sets introduced in Theorem 3.12 of Chapter 16, a characterisation which turns out to be so useful in what follows that we repeat it here for reference.

Let $(X,d)$ be a metric space and let $A \subseteq X$. A set $E \subseteq A$ is $d_A$-closed if, and only if, there is a set $F$ in $X$ that is $d$-closed and for which

$$E = A \cap F.$$

In other words, if $A \subseteq X$ is a subset in a metric space $(X,d)$ then the $d_A$-closed sets of $A$ are the intersections of $A$ with the $d$-closed sets of $X$.

## Exercise 2.1

Find a space, $(X,d)$, containing $(0,1)$ as a proper subset in which $(0,1)$ is $d$-closed.

As you may expect there is an equivalent characterisation for the open sets of a subspace.

### Theorem 2.1

Let $(X,d)$ be a metric space and let $A \subseteq X$.

A set $U \subseteq A$ is $d_A$-open if, and only if, there is a set $E$ in $X$ that is $d$-open and for which

$$U = A \cap E.$$

## Proof

The set $U$ is $d_A$-open if, and only if, the complement of $U$ in $A$, which we write as $V$, is $d_A$-closed. Therefore $U$ is $d_A$-open if, and only if, $V = \{a \in A : a \notin U\}$ is $d_A$-closed. By Theorem 3.12 of Chapter 16, we know that $V$ is $d_A$-closed if, and only if, there is a $d$-closed set $F$ in $X$ with $V = A \cap F$; see Figure 2.1. Since $F$ is $d$-closed, $X - F$ is $d$-open.

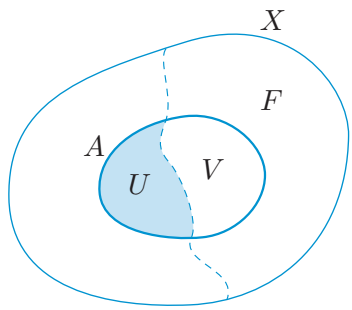



**Figure 2.1**

Now

$$\begin{aligned} A \cap (X - F) &= \{a \in A \colon a \notin F\} \\ &= \{a \in X \colon a \in A \text{ and } a \notin F\} \\ &= \{a \in X \colon a \in A \text{ and } a \in (A - F)\} \\ &= \{a \in X \colon a \in A \text{ and } a \notin V\} \\ &= \{a \in A \colon a \notin V\} \\ &= U \end{aligned}$$

and so $U$ is the intersection of $A$ with the $d$-open set $X - F$. ■

When working with subsets of subspaces, it can be necessary to carefully 'unpack' the situation so as to be very clear about exactly what is happening. (This is, in fact, true of many areas of pure mathematics.) The following worked example demonstrates this 'unpacking'.

### Worked Exercise 2.2

Let $A = (-\infty, 0) \cup (1, \infty)$ with the metric induced by $d^{(1)}$, the Euclidean metric defined on $\mathbb{R}$. Show that the subset $E \subseteq A$, given by $E = \{x \colon -7 \leq x < 0\}$ is $d_A$-closed.

#### Solution

Unpacking the definition, we are looking at $E = [-7, 0) \subseteq A \subseteq \mathbb{R}$. The space $A = (-\infty, 0) \cup (1, \infty)$ has the induced metric $d_A^{(1)}$. Therefore the $d_A^{(1)}$-closed sets in $A$ are the intersections of $A$ with the $d^{(1)}$-closed sets in $\mathbb{R}$. This means that the subset $E = [-7, 0) \subseteq A$ is $d_A^{(1)}$-closed as a set in $A$ if there is a $d^{(1)}$-closed subset $F$ of $\mathbb{R}$ such that $[-7, 0) = F \cap A$. Any interval of the form $[a, b]$ is a $d^{(1)}$-closed subset of $\mathbb{R}$, so if we find an interval $[a, b]$ with $[-7, 0) = [a, b] \cap \big((-\infty, 0) \cup (1, \infty)\big)$ then we are done. The interval $[-7, 0]$ is clearly such an interval. (In fact, any interval of the form $[-7, b]$ with $0 \leq b \leq 1$ will suffice.)

Of course you do not always have to write down the details as fully as we have in the example above. However, stopping and thinking about which space you are 'in', and exactly which metric you are using (especially if you are looking at subsets of a space with a subspace metric) is often helpful.

The next exercises give you practice in recognising $d_X^{(1)}$-closed and $d_X^{(1)}$-open subsets of $X$, where $X \subseteq \mathbb{R}$ and $(X, d_X^{(1)})$ is a metric space with the metric induced from the Euclidean metric on $\mathbb{R}$.

### Exercise 2.2

(a) Show that the subset $\{x \in \mathbb{R}: -\infty < x < 0\} = (-\infty, 0)$ is simultaneously $d$-closed and $d$-open when considered as a subset of $(X, d)$ where $X = (-\infty, 0) \cup (1, \infty)$ and $d$ is the metric induced by $d^{(1)}$, the Euclidean metric on $\mathbb{R}$.

(b) Show that the subset $\{x \in \mathbb{R}: -\infty < x \leq 0\} = (-\infty, 0]$ is simultaneously $d$-closed and $d$-open when considered as a subset of $(X, d)$ where $X = (-\infty, 0] \cup (1, \infty)$ and $d$ is the metric induced by $d^{(1)}$, the Euclidean metric on $\mathbb{R}$.

(c) Show that the subset $\{0\}$ is simultaneously $d$-closed and $d$-open when considered as a subset of $(X, d)$ where $X = \{0\} \cup (3, 4]$ and $d$ is the metric induced by $d^{(1)}$, the Euclidean metric on $\mathbb{R}$.

(d) Show that the subset $(0, 1)$ is simultaneously $d$-closed and $d$-open when considered as a subset of $\big((\mathbb{R} - \mathbb{Z}), d\big)$ where $d$ is the metric induced by $d^{(1)}$, the Euclidean metric on $\mathbb{R}$.

## Summary

We have two equivalent characterisations of continuity, which we note here for reference and comparison. We also take the opportunity to summarise some facts concerning the word 'closed'.

Let $(X, d)$ and $(Y, e)$ be metric spaces and let $f: X \to Y$ be a function.

- *The sequential characterisation of continuity*
  The function $f$ is $(d, e)$-**continuous** at $a \in X$ if, and only if, whenever $(a_k)$ is a sequence in $X$ for which $a_k \xrightarrow{d} a$ as $k \to \infty$, then the sequence $f(a_k) \xrightarrow{e} f(a)$ as $k \to \infty$.
- *The 'preimage of closed is closed' characterisation*
  The function $f$ is $(d, e)$-**continuous** if, and only if, $f^{-1}(U)$ is a $d$-closed set in $(X, d)$ whenever $U$ is an $e$-closed set in $(Y, e)$.

Let $(X, d)$ be a metric space and let $A$ be a subset of $X$.

- The set $A$ is a $d$-closed subset of $X$ if, and only if, every $d$-convergent sequence with all its terms in $A$ has its limit in $A$.
- The set $A$ is $d$-closed if, and only if, the complement, $A^c = X - A$, is $d$-open (that is, if $A^c$ contains a $d$-open ball around each of its points).

Let $(X, d)$ be a metric space. Let $E$ be a subset of $A$ where $A$ is a subset of $X$, and let $d_A$ be the metric on $A$ induced by the metric $d$ on $X$.

- The subset $E$ of $A$ is $d_A$-closed considered as a subset of $(A, d_A)$ if, and only if, there is a $d$-closed subset $F$ of $X$ with $E = A \cap F$.

# 3 Connectedness

We are now in a position to think about what we mean by 'connected'. We have an intuitive idea of what we mean for subsets in Euclidean space. We want a space such as $(0,1) \cup (2,3)$, which naturally falls into two pieces, to be disconnected. However, we want to capture the idea of 'falling apart' and so exclude cases such as $(0,1] \cup (1,2) = (0,2)$ where we have artificially 'cut' the space into two pieces. Let us examine the difference between these two cases.

Pinning down exactly what we want an intuitive idea (like 'fall apart') to mean, and translating it into the precise language of mathematics, is a very important (and often difficult) part of mathematics.

First consider $(0,1) \cup (2,3)$. This has an obvious 'gap' in the middle. Moving the second interval closer to $(0,1)$, giving $(0,1) \cup (\frac{3}{2}, \frac{5}{2})$, retains this 'gap'. Imagine moving the second interval closer and closer to the first. How small should we allow this gap to become before deciding that the space no longer 'falls apart'? Write down the space with the 'smallest' gap ....

... You may have written $(0,1) \cup (1,2)$. We want our idea of connected to encapsulate the way that $A = (0,1) \cup (1,2)$ 'falls apart' whereas $B = (0,1] \cup (1,2) = (0,2)$ does not. You may have spotted a key difference between these two cases. The set $(0,1) \cup (1,2)$ falls into two subsets, each of which is simultaneously $d_A^{(1)}$-closed and $d_A^{(1)}$-open, whereas neither $(0,1]$ nor $(1,2)$ is simultaneously $d_B^{(1)}$-closed and $d_B^{(1)}$-open. We know that for any space metric space $(X, d)$, the whole space, $X$ itself, is always simultaneously $d$-closed and $d$-open. It therefore seems that being simultaneously closed and open in this way is a rather special quality of a space which, in some sense, makes it into a 'whole piece' in its own right (rather than being merely a 'bit' of another space). We use this idea in our definition of connected.

**Definition 3.1** *Disconnection*

Let $(X, d)$ be a metric space. A **disconnection** $\{U, V\}$ of $X$ is a pair of disjoint non-empty subsets, $U$ and $V$, each of which is simultaneously $d$-closed and $d$-open, with $X = U \cup V$. We say that the space $(X, d)$ is **disconnected** if $X$ has a disconnection. It is **connected** if $X$ has no disconnection.

When we wish to emphasise the metric, we say *d-disconnection*, *d-connected* and *d-disconnected*.

Since $U$ and $V$ are required to be non-empty, neither $U$ nor $V$ can be $X$. Thus $U$ and $V$ must be proper non-empty subsets of $X$.

Recall that $A \subseteq X$ is a *proper* subset of $X$ if $A \neq X$.

### Exercise 3.1

Write down a disconnection of each of the following.

(a) The metric space $(X, d_X)$ where $X = (-\infty, 0) \cup (1, \infty)$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$.

(b) The metric space $(X, d_X)$ where $X = (-\infty, 0] \cup (1, \infty)$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$.

(c) The metric space $(X, d_X)$ where $X = \{0\} \cup (3, 4]$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$.

We now ask you to show that some sets are connected.

### Exercise 3.2

Show that the empty set is a connected space regardless of the choice of metric.

### Exercise 3.3

Let $X$ consist of a single element $X = \{x\}$. Show that $(X, d)$ is connected (where $d$ is the unique metric defined on the space and given by $d(x, x) = 0$).

The definition of a disconnection involves two subsets. However, the existence of just one non-empty, closed and open, proper subset is enough to guarantee that a space is disconnected.

### Theorem 3.2

The metric space $(X, d)$ is connected if, and only if, the only subsets of $X$ which are simultaneously $d$-closed and $d$-open are $\emptyset$ and $X$.

**Proof** By definition, we know that if $X$ has a disconnection $U \cup V$ then $U$ and $V$ are non-empty, proper subsets which are each simultaneously $d$-closed and $d$-open. Hence if the only $d$-closed and $d$-open subsets of $X$ are $\emptyset$ and $X$, then we cannot find appropriate sets $U$ and $V$ and so $(X, d)$ is connected.

Suppose now that $U$ is a non-empty and proper $d$-closed and $d$-open subset of $X$. We must show that $(X, d)$ is disconnected. Define $V = U^c$. Since $U$ is $d$-closed, $V$ is $d$-open and since $U$ is $d$-open, $V$ is $d$-closed. Since $U$ is proper, $V$ is non-empty and since $V$ is proper, $U$ is non-empty. Hence $U$

and $V$ are disjoint non-empty $d$-closed and $d$-open sets and their union is $X$. Thus $\{U, V\}$ is a disconnection of $X$, and so $(X, d)$ is disconnected. ■

## Worked Exercise 3.3 *Cantor space*

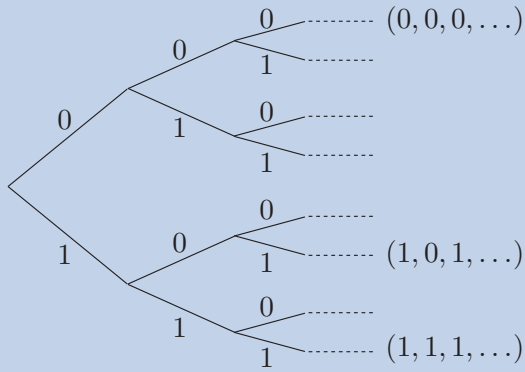



**Figure 3.1**

Cantor space, $\mathbf{C}$, and Cantor distance, $d_{\mathbf{C}}$, are defined in Book D, Chapter 15, Section 2.

Show that the Cantor space $\mathbf{C}$ is $d_{\mathbf{C}}$-disconnected.

### Solution

Recall from Chapter 15 that the Cantor space $\mathbf{C}$ consists of all infinite sequences of zeros and ones. In Chapter 16 we showed that the open balls in $\mathbf{C}$, which are $d_{\mathbf{C}}$-open, are also $d_{\mathbf{C}}$-closed sets. Therefore by Theorem 3.2 Cantor space is disconnected.

## Exercise 3.4 *The discrete metric*

Let $X$ be a set with at least two elements, and let $d_0$ be the discrete metric on $X$. Show that $(X, d_0)$ is disconnected.

Disconnectedness can also be described in terms of closed (or open) sets.

## Exercise 3.5

Prove that a metric space $(X, d)$ is disconnected if, and only if, $X$ can be expressed as the disjoint union of two non-empty $d$-closed sets.

A similar proof shows that a metric space $(X, d)$ is disconnected if, and only if, $X$ can be expressed as the disjoint union of two non-empty $d$-open sets. We have the following result.

### Proposition 3.4

The following statements are equivalent for a metric space $(X, d)$.

(a) $(X, d)$ is $d$-disconnected.

(b) $X$ can be expressed as the disjoint union of two non-empty $d$-closed sets.

(c) $X$ can be expressed as the disjoint union of two non-empty $d$-open sets.

Some texts define a metric space $(X, d)$ as being disconnected if it can be expressed as the disjoint union of two non-empty $d$-open sets.

### Worked Exercise 3.5

Determine whether $(\mathbb{Q}, d_{\mathbb{Q}})$ is connected, when $d_{\mathbb{Q}}$ is the induced metric on $\mathbb{Q}$ inherited from $\mathbb{R}$ with the Euclidean metric.

#### Solution

We look for a way to divide $\mathbb{Q}$ into two disjoint, $d_{\mathbb{Q}}$-closed subsets of $\mathbb{Q}$. The set $\mathbb{Q}$ has the metric, $d_{\mathbb{Q}}$, induced by the Euclidean metric on $\mathbb{R}$. Therefore, by Book D, Chapter 16, Theorem 3.12, we know a subset of $\mathbb{Q}$ is $d_{\mathbb{Q}}$-closed if, and only if, it is the intersection of a $d^{(1)}$-closed subset of $\mathbb{R}$ with $\mathbb{Q}$.

Let $x$ be an irrational number, such as $\sqrt{2}$. Then $E = (-\infty, \sqrt{2}] \cap \mathbb{Q}$ is the intersection of a $d^{(1)}$-closed set in $\mathbb{R}$ with $\mathbb{Q}$, and so is $d_{\mathbb{Q}}$-closed. Similarly, $F = [\sqrt{2}, \infty) \cap \mathbb{Q}$ is $d_{\mathbb{Q}}$-closed. These sets are disjoint (since $\sqrt{2} \notin \mathbb{Q}$) and $\mathbb{Q} = E \cup F$. Therefore, by Proposition 3.4, $(\mathbb{Q}, d_{\mathbb{Q}})$ is disconnected.

We ask you to try a similar exercise.

### Exercise 3.6

Determine whether $(\mathbb{R} - \mathbb{Q}, d_{\mathbb{R}-\mathbb{Q}})$ is connected, when $d_{\mathbb{R}-\mathbb{Q}}$ is the metric on $\mathbb{R} - \mathbb{Q}$ induced by the Euclidean metric on $\mathbb{R}$.

As we would expect, homeomorphic spaces come in the same number of 'pieces', and connectedness is a topological property.

### Lemma 3.6

Let $(X, d)$ and $(Y, d')$ be metric spaces, with $f\colon X \to Y$ a $(d, d')$-continuous map. If $(X, d)$ is $d$-connected then $\left(f(X), d'_{f(X)}\right)$ is a $d'_{f(X)}$-connected metric space. In other words, the continuous image of a connected metric space is connected.

**Proof** Let $(X, d)$ be a connected metric space and let $f\colon X \to Y$ be a $(d, d')$-continuous map from $X$ to $Y$. Assume for a contradiction that $E \cup F = f(X)$ is a $d'_{f(X)}$-disconnection of the image $f(X)$. Then $E$ and $F$ are each proper $d'_{f(X)}$-closed non-empty subsets of $f(X)$. Consider the preimages $E_X = f^{-1}(E)$ and $F_X = f^{-1}(F)$. Clearly $X = E_X \cup F_X$ and each of the sets $E_X$ and $F_X$ is non-empty. Also $E_X \cap F_X = f^{-1}(E) \cap f^{-1}(F) = f^{-1}(E \cap F)$. Therefore since $E$ and $F$ are disjoint we must have $E_X$ and $F_X$ disjoint. The function $f$ is $(d, d')$-continuous, so the preimage of a $d'_{f(X)}$-closed subset of $Y$ is a $d$-closed subset of $X$ (by Book D, Chapter 16, Theorem 7.1). Therefore both $E_X$ and $F_X$ are $d$-closed and so form a $d$-disconnection of $X$. This contradicts the fact that $X$ is $d$-connected, so the assumption is false and $f(X)$ is $d$-connected. ■

**Corollary 3.7**

Connectedness is preserved by homeomorphism.

**Proof** Let $(X, d)$ and $(Y, e)$ be homeomorphic metric spaces and let $f\colon X \to Y$ be a homeomorphism between them. The map $f\colon X \to Y$ is onto and a $(d, e)$-continuous map. Therefore if $X$ is $d$-connected then Lemma 3.6 implies that $Y$ must be $e$-connected. Also $f^{-1}\colon Y \to X$ is an onto $(e, d)$-continuous map from $Y$ to $X$ so, by the same lemma, $X$ is $d$-connected if $Y$ is $e$-connected. Therefore $X$ is $d$-connected if, and only if, $Y$ is $e$-connected. ■

The converse of Lemma 3.6 is not true. We ask you to show this in the next exercise.

**Exercise 3.7**

Give an example of two metric spaces $(X, d)$ and $(Y, e)$ and an onto $(d, e)$-continuous function $f\colon X \to Y$ such that $(X, d)$ is disconnected but $(Y, e)$ is connected.

## 3.1 Using a continuous function to characterise disconnectedness

We now know that the continuous image of a connected space is connected. This means that if the continuous image of a space is disconnected, then that space must be disconnected. In particular, we can characterise connectedness by considering continuous maps from a space to a disconnected space. Arguably the 'simplest' disconnected space is $\big(\{0, 1\}, d_0\big)$ where $d_0$ is the discrete metric. We use this space in our characterisation.

There is nothing special about the numbers 0 and 1: we could use any set with two elements here.

**Definition 3.8** *Characteristic function*

Let $(X, d)$ be a metric space. If $U \subseteq X$, then the **characteristic function** of $U$ is the function $\chi_U \colon X \to \{0, 1\}$ (Figure 3.2) defined by

$$\chi_U(x) = \begin{cases} 1, & \text{if } x \in U, \\ 0, & \text{if } x \in U^c. \end{cases}$$

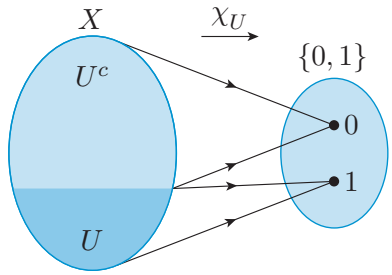



**Figure 3.2**

In other words, $\chi_U(x)$ is simply the function which takes all elements of $U$ to 1 and everything else to 0. In the following exercises we ask you to investigate what it means for this function to be continuous.

**Exercise 3.8**

Let $(X, d)$ be a metric space and let $U \subseteq X$. Show that if the characteristic function $\chi_U \colon X \to \{0, 1\}$ is $(d, d_0)$-continuous, where $d_0$ is the discrete metric, then $U$ is both $d$-closed and $d$-open.

**Exercise 3.9**

Let $(X, d)$ be a metric space and let $U \subseteq X$. Show that if $U$ is both $d$-closed and $d$-open then the characteristic function $\chi_U \colon X \to \{0, 1\}$ is $(d, d_0)$-continuous (where $d_0$ is the discrete metric on $\{0, 1\}$).

More generally, let $(X, d)$ be a metric space. Suppose that we can find a $(d, d_0)$-continuous function $f \colon X \to Y$ where $Y = \{0, 1\}$ and $d_0$ is the discrete metric. Then

$$U = f^{-1}(\{1\}) \quad \text{and} \quad U^c = f^{-1}(\{0\})$$

are disjoint closed and open sets. Thus $X$ is disconnected (unless one of $U$ or $U^c$ is the empty set, in which case $f$ is a constant function).

Here, $f$ could be one of two constant functions: $f(x) = 0$ for all $x \in X$, or $f(x) = 1$ for all $x \in X$.

This gives the following useful characterisation of connectedness.

**Theorem 3.9**

Let $(X, d)$ be a metric space. Then $X$ is $d$-connected if, and only if, every $(d, d_0)$-continuous function $f \colon X \to (Y, d_0)$, where $Y = \{0, 1\}$ and $d_0$ is the discrete metric, is constant.

### Remark

One consequence of this theorem is that if $X$ is connected, then no function from $X$ to $\{0, 1\}$ can be both continuous and *onto*.

## 3.2 Connectedness and closure

Many properties are not preserved by closure: the closure of the rational numbers is the real numbers and so countability is not preserved by closure

In this section we use the characterisation of connectedness given by Theorem 3.9 to show that connectedness is preserved by the closure operation. In other words, if $(X, d)$ is a metric space and $A \subseteq X$ is a connected subset of $X$, then $\mathrm{Cl}_{(X,d)}(A)$ is a connected subset of $X$. Intuitively, when forming the closure of a set we are simply adding in all the limit points of any sequences in the set. Therefore we would expect connectedness to be preserved since all the points that we add in are 'very close' to the set.

In fact, we prove a stronger result and show that if $A$ is connected then every subset of $\mathrm{Cl}_{(X,d)}(A)$ containing $A$ is connected. In order to do this, we first need a lemma which says that if $(X, d)$ is a metric space and $A$ is a subset of $X$ then any non-empty $d$-open subset, $V$, of $\mathrm{Cl}_{(X,d)}(A)$, the closure of $A$, must intersect $A$.

**Lemma 3.10**

Let $(X, d)$ be a metric space and let $A$ be a non-empty subset of $X$. Suppose that $V$ is a non-empty $d$-open subset of $\mathrm{Cl}_{(X,d)}(A)$. Then $A \cap V$ is non-empty.

**Proof** We suppose, by way of contradiction, that $A \cap V = \emptyset$. Now since $V$ is $d$-open, its complement $V^c$ is $d$-closed and hence $\mathrm{Cl}_{(X,d)}(A) \cap V^c$ is also $d$-closed. By our assumption, $A \subseteq V^c$ (and clearly $A \subseteq \mathrm{Cl}_{(x,d)}(A)$), so $A \subseteq \mathrm{Cl}_{(X,d)}(A) \cap V^c$. Therefore $\mathrm{Cl}_{(X,d)}(A) \cap V^c$ is a $d$-closed set containing $A$. Since, by definition, $\mathrm{Cl}_{(X,d)}(A)$ is the smallest $d$-closed set containing $A$, we have that $\mathrm{Cl}_{(X,d)}(A) \subseteq \mathrm{Cl}_{(X,d)}(A) \cap V^c$ and so $\mathrm{Cl}_{(X,d)}(A) \subseteq V^c$. That means that $V \subseteq (\mathrm{Cl}_{(X,d)}(A))^c$, which contradicts the assumption that $V$ is a non-empty $d$-open subset of $\mathrm{Cl}_{(X,d)}(A)$. Thus $A \cap V$ is non-empty as required. ■

Before moving on, we pause for a moment to unpack exactly what we mean when we talk about a *connected subset* of $X$. Let $(X, d)$ be a metric space, let $A \subseteq X$ and let $d_A$ be the metric on $A$ induced by the metric on $X$. When we talk about the connectedness of $A$, we want to consider whether we can find a closed and open subset $E$ of $A$. The question then arises as to whether we want $E$ to be $d_A$-closed and $d_A$-open, or $d$-closed and $d$-open. (For example, take $E = [1, 2)$, $A = (0, 2)$ and $X = \mathbb{R}$. Then $E$ is closed when considered as a subset of $A$, but is not closed when considered a subset of $\mathbb{R}$.) To pre-empt any confusion, we make the following definition.

**Definition 3.11** *Connected subspace*

Let $(X, d)$ be a metric space with $A \subseteq X$. Then $A$ is a **connected subspace** of $X$ if $(A, d_A)$ is connected when viewed as a metric space in its own right.

In other words, if $(X, d)$ is a metric space and $A \subseteq X$, then we look for two subsets which are each simultaneously $d_A$-closed and $d_A$-open, when considered as subsets of $A$, in order to form a disconnection of $A$.

We now consider a metric space $(X, d)$ and prove that if $A$ is a $d_A$-connected subset of $X$, and $A \subseteq B \subseteq \mathrm{Cl}_{(X,d)}(A)$ then $B$ is a $d_B$-connected subset of $X$. In other words, we want to prove that adding in any of the extra points needed to form the closure of a set cannot 'disconnect it'. We do this by contradiction. In the following proof we assume that $B$ is $d_B$-disconnected, deduce there must be a $(d_B, d_0)$-continuous *onto* function $f\colon B \to \{0, 1\}$ and show that the restriction of this function to $A$ is also onto (which is not possible since $A$ is $d_A$-connected). We use Lemma 3.10 to show that the ($d$-open) preimage under $f$ of *each* of the two points 0 and 1 must intersect $A$ (and so $f|_A$ must be onto).

$d_A$ and $d_B$ are the metrics on $A$ and $B$ induced by the metric $d$ on $X$.

**Theorem 3.12**

Let $(X, d)$ be a metric space and let $A$ be a $d_A$-connected subset of $X$. Write $\mathrm{Cl}(A)$ for $\mathrm{Cl}_{(X,d)}(A)$. Suppose that

$$A \subseteq B \subseteq \mathrm{Cl}(A).$$

Then $B$ is $d_B$-connected.

In particular, $\mathrm{Cl}(A)$ is $d_{\mathrm{Cl}(A)}$-connected.

**Proof** The proof is by contradiction. Assume that $B$ is $d_B$-disconnected. Then Theorem 3.9 implies that there is a $(d_B, d_0)$-continuous function $f$ from $B$ *onto* $\{0, 1\}$. Now the restriction of a continuous function between metric spaces is continuous (by Book D, Chapter 15, Theorem 1.6) and so $f|_A\colon A \to \{0, 1\}$, the restriction of $f$ to $A$, is $d_A$-continuous.

Since $f\colon B \to \{0, 1\}$ is onto, we can find a point $b \in B$ such that $f(b) = 0$.

Consider $V = f^{-1}(\{0\})$, the preimage of $\{0\}$. Since $\{0\}$ is $d_0$-open in $\{0, 1\}$, we must have that $V$ is $d_B$-open in $B$. Since $V$ is $d_B$-open in $B$, by Theorem 2.1 we can find a $d$-open set $V'$ with $V = V' \cap B$. Since $V'$ is $d$-open, applying Theorem 2.1 a second time gives us a $d_{\mathrm{Cl}(A)}$-open set $U = V' \cap \mathrm{Cl}(A)$.

Since $B \subseteq \mathrm{Cl}(A)$, we have $b \in V = (V' \cap B) \subseteq (V' \cap \mathrm{Cl}(A)) = U$ and so $U$ is a non-empty $d_{\mathrm{Cl}(A)}$-open subset of $\mathrm{Cl}(A)$. Here by Lemma 3.10 the set $U$ has non-empty intersection with $A$ (Figure 3.3), and we can find $a \in U \cap A$. Then $a \in U$, and so $f(a) = 0$.

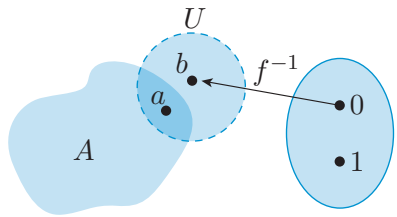



**Figure 3.3**

Since $f$ is onto, there is also a point $b' \in B$ for which $f(b') = 1$.

Repeating the same argument, we conclude that there is a point $a' \in A$ for which $f(a') = 1$. Thus $f$, when restricted to $A$, is also onto. Since $(A, d_A)$ is $d_A$-connected, this is a contradiction since no such continuous and onto function exists (by the remark following Theorem 3.9). Thus $B$ is $d_B$-connected. ■

We have shown that the continuous image of a connected set is connected and that the closure of a connected set is connected. In the next section we think a little more about the nature of the pieces into which a set can fall.

## 3.3 Components

A disconnected space can often fall apart in different ways.

**Exercise 3.10**

Write down two different disconnections of $(\mathbb{Q}, d_{\mathbb{Q}})$.

For other spaces, such as $\left((0,1)\cup(1,2), d^{(1)}\right)$, there is only one obvious disconnection. In some sense, $\mathbb{Q}$ is more granular than $\left((0,1)\cup(1,2), d^{(1)}\right)$, which comes in two distinct 'lumps'. Notice that the metric is of fundamental importance here. Putting the discrete metric onto $(0,1)\cup(1,2)$ changes its nature entirely, since with the discrete metric, $(0,1)\cup(1,2)$ becomes at least as granular as $\mathbb{Q}$.

It is natural to ask whether there is an idea in which we can talk about the *largest* 'lumps' in a space. In other words, the largest connected subsets of a space.

Let $(X,d)$ be a metric space and let $x \in X$. We know, from Exercise 3.3, that the single element set $\{x\}$ is connected. Therefore any given point $x$ in $X$ is always a member of at least one connected subset of $X$, namely $\{x\}$. In fact, it turns out that there is always a unique *largest* connected subset of $X$ containing $x$.

We will pause briefly to clarify what we mean by 'largest' here. When we describe a set $A$ as the *largest* connected subset of $X$ containing $x$, we mean that if $B$ is a connected subset of $X$, and $x \in B$, then $B \subseteq A$. In other words, $A$ is a connected subset that contains any connected subset which contains $x$.

**Definition 3.13** *Connected component*

Let $(X,d)$ be a metric space and let $x \in X$.

The **$d$-component** $C_d(x)$ of $x$ in $X$ is the largest $d$-connected subset of $X$ that contains $x$.

The existence of $C_d(x)$ is proved in Theorem 3.16.

**Definition 3.14** *Totally disconnected*

$X$ is **totally $d$-disconnected** if, for each point $x \in X$, the $d$-component of $x$ is the set $\{x\}$.

**Remark**

When we do not wish to emphasise the metric, we say that $X$ is totally disconnected.

It is not obvious that the component of a point always exists. To prove that it does, we first need a lemma in which we prove that under certain conditions the union of a family of connected subsets of a space $X$ is connected. This is not true in general (for example, $X = (1, 2) \cup (3, 4)$ is disconnected). We need to make sure that the sets are 'stuck together' in a suitable way, with no set, or collection of sets, separated from the rest.

The following lemma states that the union, $A = \bigcup_{i \in I} A_i$, of a family of connected subsets is connected if the intersection of all the sets is non-empty. The proof is by contradiction. It may seem rather technical, but the idea is straightforward. We simply assume (for a contradiction) that the set $A$ falls into two pieces, $U$ and $V$, and show that this disconnection of $A$ must cut at least one of the components, $A_j$, into two pieces $U_j = U \cap A_j$ and $V_j = V \cap A_j$ (contradicting the fact that $A_j$ is connected). The technicalities come about in proving that all the relevant sets are non-empty.

### Lemma 3.15

Let $(X, d)$ be a metric space, and let $\{A_i : i \in I\}$ be a family of subsets of $X$ and write $d_{A_i}$ for the metric on $A_i$ induced by $d$. Assume that the intersection of the family, $\bigcap_{i \in I} A_i$, is non-empty and that for each $i \in I$, $A_i$ is $d_{A_i}$-connected. Then $A = \bigcup_{i \in I} A_i$ is $d_A$-connected (where $d_A$ is the metric on $A$ induced by the metric $d$ on $X$).

**Proof** Note that all our sets are subsets of $X$ and so all inherit the induced subspace metric from $X$. We want to prove that $A$ is a $d_A$-connected subset of $X$. In other words, we want to show that $(A, d_A)$ is $d_A$-connected, where $d_A$ is the metric on $A$ induced by the metric $d$ on $X$. Therefore we work with respect to $A$ and use the metric $d_A$.

Without loss of generality, we assume that the index set $I$ has at least two elements, since otherwise there is nothing to prove. The proof is by contradiction.

Assume that $A$ is $d_A$-disconnected, and so we can find two disjoint sets, $U$ and $V$, each of which is simultaneously $d_A$-closed and $d_A$-open, with $A = U \cup V$ and $U \cap V = \emptyset$. Let $a$ belong to the non-empty intersection of all the subsets $A_i = \bigcap_{i \in I} A_i$ (Figure 3.4). Then $a \in A$ must belong to either $U$ or $V$. Without loss of generality, we may assume $a \in U$. So $a$ is in the intersection of all of our sets $A_i$ and is in $U$. Since $V$ is non-empty, at least one of the sets $A_i$ must meet $V$. Assume this is $A_j$. So the intersection $A_j \cap V$ is non-empty. Let $V_j = A_j \cap V$.
Since $a \in \bigcap_{i \in I} A_i$, we must have $a \in A_j$. Since $a \in A_j$ and $a \in U$, $A_j \cap U$ is also non-empty. Let $U_j = A_j \cap U$.

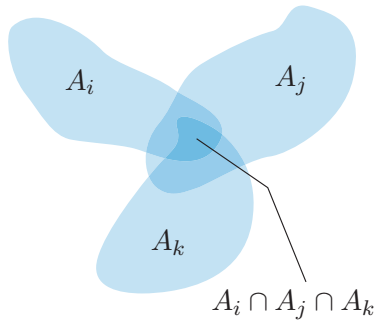



**Figure 3.4**

Consider the pair $\{U_j, V_j\}$. We have just shown them to be non-empty (by carefully selecting $A_j$ so that $V_j$ would be non-empty and by showing that $a \in U_j$). Now $U$ is $d_A$-closed in $A$ and so $U_j = U \cap A_j$ is $d_{A_j}$-closed in $A_j$. Similarly, $V_j$ is $d_{A_j}$-closed in $A_j$.

They are also disjoint since $U$ and $V$ are disjoint and $U_j \subseteq U$ and $V_j \subseteq V$. Also their union is $A_j$ since:

$$U_j \cup V_j = (A_j \cap U) \cup (A_j \cap V) = A_j \cap (U \cup V) = A_j \cap A = A_j.$$

So $A_j = U_j \cup V_j$ has been expressed as the disjoint union of two $d_{A_j}$-closed sets, and so is $d_{A_j}$-disconnected, which gives us our contradiction (as we assumed that each set $A_j$ was $d_{A_j}$-connected). Therefore $A$ is a $d_A$-connected subset of $X$. ■

We can now prove that the component of a point always exists and is well defined. Note that in the following proof the set $\mathcal{F}_x$ is a *set of sets*. An element of $\mathcal{F}_x$ is a connected *set* containing $x$. The component $C_d(x)$ is just one set, and is equal to the union of all the sets in $\mathcal{F}_x$.

**Theorem 3.16**

Let $(X, d)$ be a metric space and let $x \in X$. Let

$$\mathcal{F}_x = \{A \subseteq X : x \in A, \text{ and } A \text{ is } d_A\text{-connected}\}.$$

Then $C_d(x)$, the $d$-component of $x$, exists and is given by

$$C_d(x) = \bigcup_{A \in \mathcal{F}_x} A.$$

Note $x \in C_d(x)$ since $\{x\} \in \mathcal{F}_x$.

**Proof** The single element set $\{x\}$ is in $\mathcal{F}_x$ and so $C_d(x)$ is not empty. Also since every set in $\mathcal{F}_x$ contains $x$ (by definition), the intersection of all the sets in $\mathcal{F}_x$ is non-empty. So, by Lemma 3.15, $\bigcup_{A \in \mathcal{F}_x} A$ is connected.

If $A$ is a connected subset of $X$ that contains $x$, then $A$ is one of the sets in $\mathcal{F}_x$, and so $A$ must be part of the union of all of these sets. In other words, since $A$ is an element of $\mathcal{F}_x$, then $A$ is one of the sets in the union $\bigcup_{A \in \mathcal{F}_x} A$, which is $C_d(x)$. So $\bigcup_{A \in \mathcal{F}_x} A$ is the unique largest connected set that contains $x$. ■

Note that, if $X$ is $d$-connected, then $C_d(x) = X$ for each $x \in X$.

**Example 3.17**

The components of $X = \left((0,1) \cup (2,3), d^{(1)}_{(0,1)}\right)$ are $(0,1)$ and $(2,3)$.

We prove that all open intervals are connected in Section 4.

**Worked Exercise 3.18**

Let $\mathbb{Q}$ have the induced metric $d^{(1)}_{\mathbb{Q}}$ from $\mathbb{R}$ with the Euclidean metric. Show that the component of each point $q \in \mathbb{Q}$ is $\{q\}$.

### Solution

The proof is by contradiction. Let $q \in \mathbb{Q}$. Write $d$ for the induced metric $d_{\mathbb{Q}}^{(1)}$ on $\mathbb{Q}$ induced by the Euclidean metric on $\mathbb{R}$. Assume, for a contradiction, that $C_d(q)$ consists of more than one point, so there is $p \neq q$ with $p \in C_d(q)$. Then, by an argument similar to that used in the solution to Worked Exercise 3.9 in Chapter 16, there is an irrational number $x$ such that $\min\{p, q\} < x < \max\{p, q\}$. Set

$$U = \{r \in C_d(q) : r < x\} \quad \text{and} \quad V = \{r \in C_d(q) : r > x\}.$$

Then $U$ and $V$ are non-empty and disjoint, and their union is $C_d(q)$. Now $(-\infty, x) \cap \mathbb{Q}$ is $d$-open in $\mathbb{Q}$, and so $U = ((-\infty, x) \cap \mathbb{Q}) \cap C_d(q)$ is $d_{C_d(q)}$-open (where $d_{C_d(q)}$ is the metric on $C_d(q)$ induced by the metric $d$ on $\mathbb{Q}$).

Similarly, $V$ is $d_{C_d(q)}$-open for the induced metric on $C_d(q)$. Thus $\{U, V\}$ is a disconnection of $C_d(q)$. But this is impossible, since $C_d(q)$ is connected. This contradiction shows that there can be no point $p \neq q$ in $C_d(q)$. Thus, $C_d(q) = \{q\}$.

Since $\mathbb{Q}$ consists only of rational points, $x \notin C_d(q)$.

Do not be put off by the 'nested' induced metrics in the above example. We start with the Euclidean metric, $d^{(1)}$ on $\mathbb{R}$, and use this to induce a metric $d_{\mathbb{Q}}^{(1)}$ on $\mathbb{Q}$. We then take a subset $C_d(q)$ of $\mathbb{Q}$ and use $d_{\mathbb{Q}}^{(1)}$ to induce a metric on this subset of $\mathbb{Q}$ (which we write as $d_{C_d(q)}$).

### Exercise 3.11

Describe the components of $X = \mathbb{R} - \mathbb{Z}$ with respect to the induced Euclidean metric. You may assume that $I = (a, b) \subseteq \mathbb{R}$ is $d_I^{(1)}$-connected. (We prove this in Section 4.)

### Exercise 3.12

Describe the components of $(0, 1)$ with the discrete metric.

The components of a space are not (necessarily) the same as the sets forming a disconnection. For example, the sets $(-\infty, 1/2] \cap \mathbb{Q}$ and $(1/2, \infty) \cap \mathbb{Q}$ form a disconnection of $\mathbb{Q}$, as do the sets $(-\infty, \sqrt{3}) \cap \mathbb{Q}$ and $(\sqrt{3}, \infty) \cap \mathbb{Q}$, whereas the components of $\mathbb{Q}$ are the singleton sets $\{q\}$ for each $q \in \mathbb{Q}$. In particular, the components of a space need not be open although, as we see later, the components are always closed.

### Worked Exercise 3.19

Show that the sets $\{q\}$, where $q \in \mathbb{Q}$, are not $d_{\mathbb{Q}}^{(1)}$-open in $\mathbb{Q}$ (where $d_{\mathbb{Q}}^{(1)}$ is the metric on $\mathbb{Q}$ induced by the Euclidean metric on $\mathbb{R}$).

### Solution

Book D, Chapter 16, Theorem 2.9 states that $A \subseteq X$, where $(X, d)$ is a metric space, is $d$-open if, and only if, we can find an $r > 0$ with $B_d(a, r) \subseteq A$ for all $a \in A$. Assume for a contradiction that $\{q\}$ is $d_{\mathbb{Q}}^{(1)}$-open in $\mathbb{Q}$. Clearly $q \in \{q\}$ so there must be an $r$ with $B_{d_{\mathbb{Q}}^{(1)}}(q, r) \subseteq \{q\}$. In other words, we must have $B_{d_{\mathbb{Q}}^{(1)}}(q, r) = \{q\}$. Choose $n > \frac{1}{r}$. Then the rational number $q + \frac{1}{n} \in B_{d_{\mathbb{Q}}^{(1)}}(q, r)$, and so $B_d(q, r) \not\subseteq \{q\}$ and we have a contradiction.

Whilst a disconnected space can often be disconnected in many different ways, there is only one way to split it into its components. In fact, a set $X$ is always the union of mutually disjoint components. We now prove this. In the following proof we define a relation, $\sim$, on a space $X$ by letting $x \sim y$ if $x$ and $y$ are in the same component. We show that $\sim$ is an equivalence relation and so must partition the set into equivalence classes. These equivalence classes make up the individual components.

An equivalence relation always partitions a set into disjoint equivalence classes.

### Theorem 3.20

Let $(X, d)$ be a metric space. Let us write $x \sim y$ to mean that $y \in C_d(x)$. Then $\sim$ is an equivalence relation on $X$. Consequently, $X$ is the union of its mutually disjoint $d$-components.

**Proof** In order to prove that $\sim$ is an equivalence relation we must show that $\sim$ is:

E1 Reflexive. For each $x \in X$, $x \sim x$.

E2 Symmetric. For each $x, y \in X$, if $x \sim y$ then $y \sim x$.

E3 Transitive. For each $x, y, z \in X$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

*Reflexive* Since $x \in C_d(x)$, it follows immediately that $x \sim x$ for each $x \in X$.

*Symmetric* Suppose $x \sim y$ – that is, $y$ is in $C_d(x)$, the component containing $x$ (Figure 3.5). Now $C_d(x)$ is connected (by definition) and contains $y$ and so must be contained in the largest connected subset of $X$ containing $y$, which is $C_d(y)$ and so $x \in C_d(y)$. Therefore $\sim$ is symmetric.

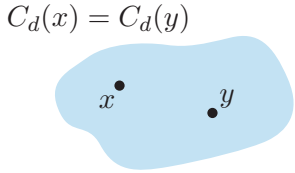



**Figure 3.5** $x \sim y \Rightarrow y \sim x$

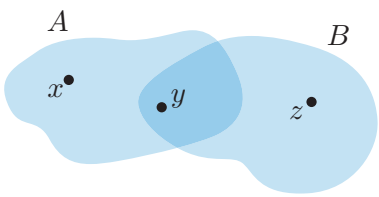



**Figure 3.6** $x \sim y,\ y \sim z$

*Transitive* Suppose $x \sim y$ and $y \sim z$. Let $A$ be a connected subset of $X$ containing $x$ and $y$, and $B$ be a connected subset of $X$ containing $y$ and $z$ (Figure 3.6). Note that $A \cap B$ contains $y$, and so is non-empty. Therefore, by Lemma 3.15, $A \cup B$ is connected. Therefore since $x$ and $z$ are in $A \cup B$, we have a connected set containing $x$ and $z$ and so it follows that $z \in C_d(x)$, and so $x \sim z$.

Hence $\sim$ is an equivalence relation, and so $X$ can be written as the disjoint union of the equivalence classes defined by $\sim$. But, by definition of $\sim$, the equivalence class of a point $x$ consists of all the points in $C_d(x)$ – namely $C_d(x)$ itself – and so each equivalence class is a component of $X$. ■

In particular, we conclude from Worked Exercise 3.18 that $\mathbb{Q}$ consists of countably many components, each of which is a singleton: $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}}\{q\}$.

There is a useful immediate corollary to this theorem.

**Corollary 3.21**

Let $(X, d)$ be a metric space. For each $y \in C_d(x)$, $C_d(y) = C_d(x)$. In other words, the component of each point in $C_d(x)$ is $C_d(x)$ itself.

## 3.4 The components are closed

From our investigation of $\mathbb{Q}$ in Worked Exercise 3.19 we know that the components of a space are not always open. In this section we show that they are always closed.

At the beginning of Section 3 we considered what happened when we moved the two intervals $(0, 1)$ and $(1, 2)$ closer and closer together. We described $X = (0, 1) \cup (1, 2)$ as disconnected and $(0, 1] \cup (1, 2) = (0, 2)$ as connected. We now consider the analogous situation in the plane.

Consider two discs of radius 1 in the plane, one centred at $(1, 1)$ and the second at $(3, 1)$ (see Figure 3.7).

First imagine that both the discs are open. In this case they do not meet at all, and the union is disconnected (as it is the disjoint union of the two open discs). If both are closed, then we expect their union to be connected (as in this case the discs meet in the point $(2, 1)$). So far, so good – but what if one is open and the other is closed? Does this third case describe a connected or disconnected subspace of the plane?

We can answer this question using the following lemma.

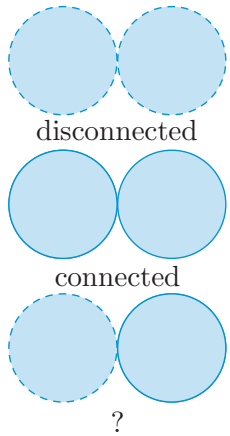



**Figure 3.7**

**Lemma 3.22**

Let $(X, d)$ be a metric space.

(a) If $\{U, V\}$ is a disconnection of a metric space $(X, d)$, then

$$\mathrm{Cl}_{(X,d)}(U) \cap V = \emptyset \text{ and } U \cap \mathrm{Cl}_{(X,d)}(V) = \emptyset.$$

(b) If there is a pair of non-empty sets $U$ and $V$ for which

$$U \cup V = X,\ \mathrm{Cl}_{(X,d)}(U) \cap V = \emptyset \text{ and } U \cap \mathrm{Cl}_{(X,d)}(V) = \emptyset,$$

then $U$ and $V$ are $d$-closed and form a disconnection of $X$.

In some texts, sets $A_1$ and $A_2$ for which $\mathrm{Cl}(A_i) \cap A_j = \emptyset$ $(i \neq j)$ are described as 'separated sets'.

**Proof** Part (a) follows from the definition of a disconnection since if $\{U, V\}$ is a disconnection of a metric space $(X, d)$ then we know from Proposition 3.4 that both $U$ and $V$ are $d$-closed. Hence $\mathrm{Cl}_{(X,d)}(U) = U$ and $\mathrm{Cl}_{(X,d)}(V) = V$, and so $\mathrm{Cl}_{(X,d)}(U) \cap V = U \cap V = \emptyset$ and $U \cap \mathrm{Cl}_{(X,d)}(V) = U \cap V = \emptyset$.

For (b) suppose $U$ and $V$ are non-empty subsets of $X$, with

$$U \cup V = X, \ \mathrm{Cl}_{(X,d)}(U) \cap V = \emptyset \text{ and } U \cap \mathrm{Cl}_{(X,d)}(V) = \emptyset.$$

To show that $U$ and $V$ form a disconnection of $X$, it is sufficient to show that they are disjoint and $d$-closed (by Proposition 3.4). Since $U \subseteq \mathrm{Cl}_{(X,d)}(U)$, we must have $U \cap V = \emptyset$, so $U$ and $V$ are disjoint. Suppose that $U$ is not $d$-closed. Since its complement in $X$ is $V$, there are points in its $d$-closure that belong to $V$. But that contradicts the hypothesis that $\mathrm{Cl}_{(X,d)}(U)$ is disjoint from $V$; hence $U$ is $d$-closed. In a similar way, we deduce that $V$ is $d$-closed. Thus $U$ and $V$ form a disconnection of $X$. ■

Returning to our disc question – if $X$ consists of two just touching discs in the plane, one $d^{(2)}$-open and one $d^{(2)}$-closed (where $d^{(2)}$ is the Euclidean metric on the plane), then the $d^{(2)}$-closure of the $d^{(2)}$-open disc meets the $d^{(2)}$-closed disc. Therefore the discs do not form a disconnection of $X$ by Lemma 3.22. In fact, $X$ is $d_X^{(2)}$-connected. Intuitively you may find it helpful to think of the open and closed discs as touching, but not meeting; whereas the two closed discs meet and the two open discs do not 'touch'.

We end this section by proving, as promised, that the components of a metric space $(X, d)$ are $d$-closed.

This result, when combined with Theorem 3.20, tells us that a metric space can be split up into disjoint closed pieces, each of which is connected. These results are very useful when it comes to identifying the components of a set.

### Lemma 3.23

Let $(X, d)$ be a metric space and let $x \in X$. Then $C_d(x)$, the component of $x$ in $(X, d)$, is $d$-closed.

**Proof** This is Exercise 3.13. ■

### Exercise 3.13

Prove Lemma 3.23.

*Hint*: you may find Theorem 3.12 useful.

## 3.5 Totally disconnected spaces

Totally disconnected spaces are spaces for which the components consist of single points (Definition 3.14). In this section we investigate some of these spaces. We have already met some such spaces. We saw in Exercise 3.12 that $\big((0,1), d_0\big)$ is totally $d_0$-disconnected and in Worked Exercise 3.18 we saw that $(\mathbb{Q}, d_{\mathbb{Q}}^{(1)})$ is totally disconnected.

### Remarks

Although $\big((0,1), d_0\big)$ and $(\mathbb{Q}, d_{\mathbb{Q}}^{(1)})$ are both totally disconnected, they are very different as metric spaces. In $(\mathbb{Q}, d_{\mathbb{Q}}^{(1)})$ the distance between two points varies: $\frac{1}{2}$ is 'closer' to $\frac{24}{50}$ than it is to $\frac{1}{50}$; whereas the distance between $\frac{1}{2}$ and $\frac{24}{50}$ is the same as the distance between $\frac{1}{2}$ and $\frac{1}{50}$ in $\big((0,1), d_0\big)$. In $(\mathbb{Q}, d_{\mathbb{Q}}^{(1)})$ there is a very well defined sense of degree of 'closeness' which we do not have in any discrete space.

In addition, $\mathbb{Q}$ is countable and $(0,1)$ is uncountable.

We now investigate the middle-third Cantor set and show that this is another example of a totally disconnected space.

### Worked Exercise 3.24

Recall that the middle-third Cantor set, $K_C$, is equal to $K_C = \bigcap_{n=1}^{\infty} K_n$ where:

$$K_0 = [0,1]$$
$$K_1 = \left[0, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, 1\right]$$
$$K_2 = \left[0, \tfrac{1}{9}\right] \cup \left[\tfrac{2}{9}, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, \tfrac{7}{9}\right] \cup \left[\tfrac{8}{9}, 1\right]$$
$$\vdots$$

with $K_n$ formed by removing the open middle third of each interval in $K_{n-1}$.

Show that the middle-third Cantor set, $K_C$, with the metric induced by the Euclidean metric on $\mathbb{R}$, is totally disconnected.

See Book D, Chapter 16, Example 1.19.

#### Solution

We need to show that each component consists of a single element. Assume that $x \neq y$ are two elements belonging to the same component, $C$, of $K_C$. Without loss of generality, assume $x < y$. From the discussion of Example 1.19 in Chapter 16, there is a point $z \notin K_C$, and so certainly not in $C$, with $x < z < y$. Therefore the sets $(-\infty, z] \cap C$ and $[z, \infty) \cap C$ form a disconnection of $C$, which is a contradiction. Therefore our assumption that $x \neq y$ must be false.

# 4 Connectedness in Euclidean space

In this section, as promised, we consider which sets in $\mathbb{R}^n$ are connected. In what follows, all sets have metrics induced by the Euclidean metric unless otherwise stated.

Our first objective is to show that $\mathbb{R}$ is itself connected. This may seem intuitively obvious, but the proof is subtle. We need to prove that it is impossible to divide $\mathbb{R}$ into two sets, each of which is simultaneously open and closed. We know that we can do this easily with $\mathbb{Q}$, so the proof will need to use a property of $\mathbb{R}$ that $\mathbb{Q}$ does not possess. The required crucial property of $\mathbb{R}$ (that $\mathbb{Q}$ does not have) is the *least upper bound property*. We state it here for reference.

We often write $\sup A = M$.

If $A \subseteq \mathbb{R}$ is a non-empty set that is bounded above, then $A$ has a **least upper bound**, or **supremum** – a real number $M$ such that:

(a) $a \leq M$, for all $a \in A$

(b) if $M' < M$, then there is some $a \in A$ with $M' < a$.

We often write $\inf A = N$.

Similarly, if $A \subseteq \mathbb{R}$ is a non-empty set that is bounded below, then $A$ has a **greatest lower bound**, or **infimum** – a real number $N$ such that:

(a) $a \geq N$, for all $a \in A$

(b) if $N' > N$, then there is some $a \in A$ with $N' > a$.

### Remark

Note that in the definition of supremum above, part (a) says that $M$ is *an* upper bound for $A$ and part (b) says that $M$ is the *least* upper bound.

Remember that $\sqrt{2} \notin \mathbb{Q}$!

For example, the set $\{x \in \mathbb{R} : x^2 < 2\}$ has a least upper bound, namely $\sqrt{2}$ in $\mathbb{R}$; but $\{x \in \mathbb{Q} : x^2 < 2\}$ does not have a least upper bound in $\mathbb{Q}$.

### Exercise 4.1

Use the least upper bound property to show that if a non-empty set $A \subseteq \mathbb{R}$ is bounded below, then the infimum exists.
*Hint*: consider the set $B = \{-a : a \in A\}$.

### Remark

Note that a closed and bounded interval, $[a, b]$, contains its supremum and infimum.

We will use the least upper bound property to show that $\mathbb{R}$ is $d^{(1)}$-connected. First we prove a lemma.

**Lemma 4.1**

Let $(X, d)$ be a metric space. A $d$-closed subset of $X$ contains all of its $d$-closure points.

A point $x$ is a closure point of a set $A$ if every open ball centred on $x$ meets the set $A$.

**Proof** Let $A$ be a $d$-closed subset of $X$. Then $A = \mathrm{Cl}_{(X,d)}(A)$ by Theorem 3.3 of Chapter 16 and so, by Theorem 3.8 of Chapter 16, coincides with its set of closure points. ■

**Theorem 4.2**

$\mathbb{R}$ with the Euclidean metric is connected.

**Proof** The proof is by contradiction. We work with respect to the Euclidean metric, $d^{(1)}$. Suppose that $\mathbb{R}$ is disconnected, and let $\{U, V\}$ be a disconnection. Thus $U$ and $V$ are disjoint non-empty $d^{(1)}$-closed and $d^{(1)}$-open subsets of $\mathbb{R}$ with union $\mathbb{R}$.

Consider the pair of intervals $(-\infty, 0]$ and $[0, \infty)$. These are both $d^{(1)}$-closed and their union is $\mathbb{R}$, but they are not disjoint – they have the point 0 in common.

Since $U \cup V = \mathbb{R}$, the point 0 lies in one of the sets $U$ and $V$. Without loss of generality, we can suppose that $0 \in V$. Since $U \neq \emptyset$, at least one of $(-\infty, 0] \cap U$ and $[0, \infty) \cap U$ is non-empty.

*Case 1:* $(-\infty, 0] \cap U \neq \emptyset$

Let $A = (-\infty, 0] \cap U$.

Since $A$ is the intersection of two $d^{(1)}$-closed sets, it is $d^{(1)}$-closed. It is also bounded above by 0.

Since $0 \in V$ and $U \cap V = \emptyset$, $0 \notin U$, and so 0 cannot be in $A$. Hence

$$A = (-\infty, 0) \cap U.$$

In other words, removing the point 0 does not change $A$. As $U$ is $d^{(1)}$-open, we now have $A$ expressed as the intersection of two $d^{(1)}$-open sets and so $A$ is $d^{(1)}$-open.

Hence $A$ is $d^{(1)}$-closed and $d^{(1)}$-open, non-empty and bounded above (by 0). We now use these four properties of $A$ to reach a contradiction to our assumption that $(-\infty, 0] \cap U \neq \emptyset$.

By the least upper bound property, $A$ has a least upper bound – we denote this by $M$. We first show that $M$ is a $d^{(1)}$-closure point of $A$.

Consider a $d^{(1)}$-open ball centred on $M$, $B(M, r)$ for an arbitrary $r > 0$. This contains $M - \frac{r}{2}$. Since $M - \frac{r}{2} < M$ we must be able to find an $a \in A$ with $M \geq a > M - \frac{r}{2}$ (because $M$ is a *least* upper bound for $A$). Since $M \geq a > M - \frac{r}{2}$ we clearly have $a \in B(M, r)$.

Note that $a$ may be equal to $M$.

Therefore $a \in A \cap B(M, r)$ and so $A \cap B(M, r)$ is non-empty. Therefore, since $r$ was arbitrary, $M$ is a $d^{(1)}$-closure point of $A$ and so, by Lemma 4.1 $M$ is an element of $A$.

Since $A$ is also $d^{(1)}$-open, it contains a $d^{(1)}$-open ball centred at each of its points. In particular, it must contain a ball centred at $M$. So there is a number $r > 0$ such that $B(M, r) = (M - r, M + r) \subseteq A$.

But then $A$ contains points $y$ with $M < y < M + r$ – for example, $M + \frac{1}{2}r$ – and so $M$ cannot be an upper bound for $A$.

We have arrived at a contradiction: so Case 1 does not apply.

*Case 2:* $[0, \infty) \cap U \neq \emptyset$

In this case, we let $A = [0, \infty) \cap U$, and use the greatest lower bound property to obtain a contradiction in a similar way: so Case 2 does not apply.

We conclude that $\mathbb{R}$ is connected. ■

Intuition suggests that every interval is connected. We ask you to prove this for the interval $(0, 1)$.

### Exercise 4.2

Show that the interval $(0, 1)$ is a connected subset of $\mathbb{R}$ (with the Euclidean metric).

### Exercise 4.3

Describe the components of $(0, 1)$ with the Euclidean metric.

Note that in Exercise 3.12 we showed that the components of $\big((0, 1), d_0\big)$ were the singleton sets $\{x\}$ for each $x \in (0, 1)$, whereas here we have shown that $\big((0, 1), d^{(1)}_{(0,1)}\big)$ has just one component. This highlights just how important the metric is. The space $\big((0, 1), d^{(1)}_{(0,1)}\big)$ comes in a single 'lump', whereas $\big((0, 1), d_0\big)$ falls into infinitely many tiny pieces!

It is now a simple step to show that any open interval is connected.

### Proposition 4.3

Let $a < b$ be real numbers. The interval $(a, b)$ is a connected subset of $\mathbb{R}$ (with the Euclidean metric).

**Proof** From the discussion at the beginning of Subsection 1.1, we know that any two open intervals of $\mathbb{R}$ are homeomorphic. As connectedness is preserved by homeomorphism, Exercise 4.2 shows that any open interval of $\mathbb{R}$ is connected. ■

### Exercise 4.4

Let $a < b$ be real numbers. Show that the intervals $[a, b)$, $(a, b]$ and $[a, b]$ are connected subsets of $\mathbb{R}$ (with the Euclidean metric).

We have shown that all bounded intervals are connected. What about unbounded intervals such as $(a, \infty)$? The continuous image of a connected space is connected, so can we find a continuous map from $\mathbb{R}$ to $(a, \infty)$? It turns out that we can. The exponential map takes $\mathbb{R}$ to $(0, \infty)$ so we use this and rescale to get the map $\phi\colon \mathbb{R} \to (a, \infty)$ given by $\phi(x) = a + e^x$. This is continuous, showing that $(a, \infty)$ is connected for each $a \in \mathbb{R}$. Likewise we can use the map $\psi(x) = a - e^x$ to show that $(-\infty, a)$ is connected. Theorem 3.12 now shows that $(-\infty, a]$ and $[a, \infty)$ are also connected.

Certainly $\emptyset$ and single-point subsets of $\mathbb{R}$ are connected and so we have shown that all the intervals in $\mathbb{R}$ are connected. In fact, it turns out that the *only* connected subsets of $\mathbb{R}$ are intervals.

### Theorem 4.4

A subset $A$ of $\mathbb{R}$ is connected for the Euclidean metric if, and only if, it is an interval.

**Proof** We have already shown that all intervals are connected. It remains to show that if $A$ is connected, then it is an interval. Informally, if $A$ is not an interval then it must have at least one 'hole' in it. In this proof by contradiction we use this hole to form a disconnection of $A$.

Let $A$ be connected, and suppose that $A$ is *not* an interval. Now the empty set and single-point sets are intervals, so $A$ must contain at least two points. Since $A$ is not an interval, there must be real numbers $a, b, c$ with $a, b \in A$ and $a < c < b$ such that $c \notin A$ – otherwise $A$ would be an interval. Let $\widetilde{E} = (-\infty, c]$ and $\widetilde{F} = [c, \infty)$. We now show that $E = \widetilde{E} \cap A$ and $F = \widetilde{F} \cap A$ form a disconnection of $A$.

The sets $E$ and $F$ are $d_A$-closed for the metric induced on $A$ by the Euclidean metric on $\mathbb{R}$ since they are intersections with $A$ of the $d^{(1)}$-closed sets $\widetilde{E}$ and $\widetilde{F}$. Neither is empty since $a \in \widetilde{E}$ and $b \in \widetilde{F}$. The intersection of $E$ and $F$ is

$$E \cap F = (\widetilde{E} \cap \widetilde{F}) \cap A = \emptyset,$$

and their union is

$$E \cup F = (\widetilde{E} \cup \widetilde{F}) \cap A = \mathbb{R} \cap A = A.$$

So we have two disjoint $d_A$-closed sets whose union is $A$. Therefore $A$ is disconnected by Proposition 3.4.

This contradicts the fact that $A$ is connected. We conclude that a connected subset of $\mathbb{R}$ must be an interval. ■

### Exercise 4.5

Write down a disconnection of $\mathbb{R} - \{a\}$ (with the metric induced by the Euclidean metric on $\mathbb{R}$), for each $a \in \mathbb{R}$.

As connectedness is a topological invariant, we have the following corollary.

### Corollary 4.5

$\mathbb{R} - \{a\}$ is not homeomorphic to $\mathbb{R}$ for the Euclidean metrics, for each $a \in \mathbb{R}$.

In fact, it follows from Theorem 4.4 that any subset of $\mathbb{R}$ that is not an interval must be disconnected.

The fact that intervals are connected can be used to show that many other sets are connected.

### Exercise 4.6

Show that the unit circle $C = \{(x, y) : x^2 + y^2 = 1\}$ is connected.

It is now a short step to the promised generalisation of the Intermediate Value Theorem from Book D, Chapter 13.

### Theorem 4.6 *Intermediate Value Theorem*

Let $(X, d)$ be a connected metric space, and let $f : X \to \mathbb{R}$ be a $(d, d^{(1)})$-continuous function. Let $a$ and $b$ be points of $X$. Then $f$ takes each value between $f(a)$ and $f(b)$.

**Proof** Since $f$ is $(d, d^{(1)})$-continuous and $X$ is connected, Lemma 3.6 implies that $f(X)$ is a connected subspace of $\mathbb{R}$, and so is an interval by Theorem 4.4. This interval contains $f(a)$ and $f(b)$, hence $f(X)$ contains each value between $f(a)$ and $f(b)$. ■

#### Remark

This remarkable theorem is seen to be an immediate consequence of three things: the domain is connected, the only connected subsets of $(\mathbb{R}, d^{(1)})$ are the intervals, and continuous images of connected sets are connected.

## 4.1 Homeomorphic subsets of $\mathbb{R}$: part 2

We can now conclude our investigation into which subsets of $\mathbb{R}$ are homeomorphic to each other. Recall that we had the following classes of intervals:

- the empty set $\emptyset$
- the singleton sets $\{a\} = [a, a]$
- the open intervals $(a, b)$ for $a < b$, $(a, \infty)$ and $(-\infty, a)$ for $a \in \mathbb{R}$, and $\mathbb{R}$
- the closed intervals $[a, b]$ for $a < b$
- the intervals of the form $[a, b)$ and $(a, b]$ for $a < b$ and $[a, \infty)$ and $(-\infty, a]$.

Intervals in the same class are homeomorphic. We now show that intervals in different classes are not homeomorphic.

### Worked Exercise 4.7

Show that $A = (0, 1)$ is not homeomorphic to $B = (0, 1]$ (when both are considered as metric spaces with the metrics induced by the Euclidean metric on $\mathbb{R}$).

#### Solution

The proof is by contradiction. We know that $A$ and $B$ are both connected. Suppose that there is a homeomorphism $\phi\colon A \to B$. Then there must be a point in $A$ which maps to the point $1 \in B$. Let $a \in A$ be this point (so $\phi(a) = 1$). Consider what happens when we take the point $a$ away from $(0, 1)$, and also remove its image, the point 1, from $(0, 1]$. The restriction rule for homeomorphisms (Proposition 1.2), implies that $A - \{a\}$ is homeomorphic to $\phi\big(A - \{a\}\big) = B - \{1\} = A$. This is a contradiction, since $A - \{a\}$ is disconnected while $A$ is connected. Thus $A$ is not homeomorphic to $B$.

A disconnection is $\{(0, a), (a, 1)\}$.

Since $(0, 1)$ is homeomorphic to $(a, b)$ and $(0, 1]$ is homeomorphic to $(a, b]$ and to $[c, d)$ (where all spaces have metrics induced by the Euclidean metric on $\mathbb{R}$), it follows from Worked Exercise 4.7 that, in general, an open interval and a half-open interval are not homeomorphic (when both are considered as metric spaces with the metric induced by the Euclidean metric on $\mathbb{R}$).

### Exercise 4.7

Show that $A = (0, 1]$ is not homeomorphic to $B = [0, 1]$ (when both are considered as metric spaces with the metric induced by the Euclidean metric on $\mathbb{R}$).

It follows that, in general, $(a, b]$ and $[a, b]$ are not homeomorphic (when both are considered as metric spaces with the metric induced by the Euclidean metric on $\mathbb{R}$).

### Remarks

1. Intuitively we would expect different types of interval to be different since the endpoints 'look' different. Imagine that you are a one-dimensional creature living inside the interval $[0, 1)$. The left-hand side would have a hard edge which you could bump right into. However, you could never reach the right-hand boundary. The same is true of the left- and right-hand edges of $[0, \infty)$, so it makes intuitive sense for the two intervals to be homeomorphic.
2. Whilst the intervals $[a, b)$ and $[a, \infty)$ share many topological properties, they do not, of course, share all properties: for instance, $[a, b)$ is bounded whilst $[a, \infty)$ is not. Homeomorphism allows us to stretch space as much as we like (and so, as we have seen, a bounded interval can be homeomorphic to an unbounded one).

## 4.2 Products of connected spaces

After this successful investigation of the connectedness of the real line, we turn our attention to the connectedness (or otherwise) of Euclidean space, $\mathbb{R}^n$ with $n > 1$.

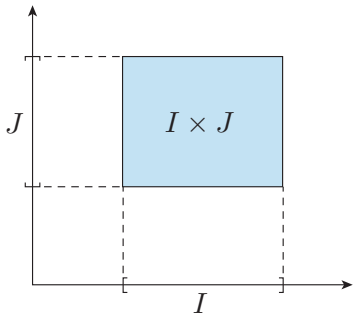



**Figure 4.1**

The fact that an $n$-fold product of connected spaces is also connected then follows by induction.

We saw in Book D, Chapters 13 and 14 that many of the properties of $\mathbb{R}^n$ are immediate consequences of the fact that $\mathbb{R}^n$ is formed by repeatedly taking products of $\mathbb{R}$. In particular, the Euclidean metric on $\mathbb{R}^2$ is a product metric of the Euclidean metric on $\mathbb{R}$ with itself.

It seems plausible that connectedness should be preserved when we take products – for example, the product of an interval with an interval gives a rectangle in the plane (Figure 4.1), a set that we expect to be connected.

In order to show that $\mathbb{R}^2$ (or a rectangle) is connected, we prove a more general result: that the product of any two connected metric spaces is itself a connected metric space.

**Theorem 4.8**

Let $(X, d_1)$ and $(Y, d_2)$ be non-empty metric spaces. The product space $(X \times Y, e)$, where $e$ is a product metric, is $e$-connected if, and only if, $(X, d_1)$ and $(Y, d_2)$ are both connected (for their respective metrics).

**Proof** We prove that if $X$ and $Y$ are connected for their respective metrics, then $X \times Y$ is $e$-connected. We leave the converse as an exercise for you (Exercise 4.8).

Suppose that $(X, d_1)$ and $(Y, d_2)$ are connected (for their respective metrics). We must show that $(X \times Y, e)$ is $e$-connected. To do this, it is enough to show that, for a given $z \in X \times Y$, the component of $z$ is the whole space $X \times Y$.

So suppose that $z = (x, y) \in X \times Y$, and let $C_e(z)$ denote the component of $z$ in $X \times Y$. In order to show that $C_e(z) = X \times Y$, we must show that if $z' \in X \times Y$ is any other point in $X \times Y$ then $z' \in C_e(z)$.

So suppose that $z' = (x', y') \in X \times Y$. We show that $z' \in C_e(z)$ by finding a connected subset $A$ of $X \times Y$ that contains $z$ and $z'$. As the component containing $(x, y)$, by definition, contains all other connected subsets of $X \times Y$, then we are done. We define $A = (X \times \{y\}) \cup (\{x'\} \times Y)$, as illustrated in Figure 4.2.

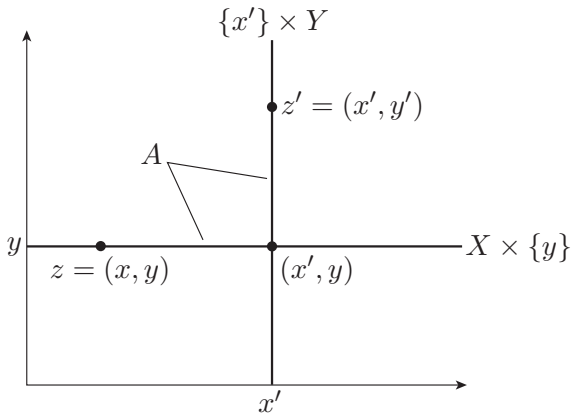



**Figure 4.2** $A = (X \times \{y\}) \cup (\{x'\} \times Y)$

Clearly, both $z$ and $z'$ are elements of $A$, so it remains only to show that $A$ is $e$-connected (as a subset of $X \times Y$).

Now $X \times \{y\}$ (as a subspace) is a $(d_1, e_{X \times \{y\}})$-continuous image of $X$ and $\{x'\} \times Y$ is a $(d_2, e_{\{x'\} \times Y})$-continuous image of $Y$. Thus, by Lemma 3.6, both $X \times \{y\}$ and $\{x'\} \times Y$ are connected.

We do not include the proofs of these straightforward but technical results.

However, the point $(x', y)$ is in both $X \times \{y\}$ and $\{x'\} \times Y$. Thus, by Lemma 3.15, $A$ is connected. ■

## Exercise 4.8

Prove the second part of this theorem, that if $(X, d_X)$ and $(Y, d_Y)$ are non-empty metric spaces and $(X \times Y, e)$ is $e$-connected, where $e$ is a product metric, then $(X, d_X)$ and $(Y, d_Y)$ are connected (for their respective metrics). *Hint*: if $(X, d_X)$ and $(Y, d_Y)$ are metric spaces, and $e$ is a product metric for $X \times Y$, then the projection function $p_X : X \times Y \to X$ is $(e, d_X)$-continuous and the projection function $p_Y : X \times Y \to Y$ is $(e, d_Y)$-continuous.

See Book D, Chapter 15, Theorem 3.11.

We have two immediate corollaries.

### Corollary 4.9

Euclidean space, $(\mathbb{R}^n, d^{(n)})$, is connected.

### Corollary 4.10

The $n$-dimensional 'rectangles' $I_1 \times \cdots \times I_n$ are connected for the Euclidean metric, where $I_i$ is an interval in $\mathbb{R}$.

### Exercise 4.9

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the continuous polynomial defined by

$$f(x_1, x_2) = 5x_1^5 + 5x_1^3x_2^2 - x_2^5.$$

Show that there is a point in the plane at which $f$ attains the value 42.

We now know enough to deduce that many other spaces are connected. The following discussion is informal but should give you an idea of how the above results can be applied.

We know that the unit circle $S$ is a connected subset of the plane from Exercise 4.6. What happens if we take a product of a circle with itself? Imagine taking a flexible cylindrical pipe (such as a section of particularly bendy hosepipe).

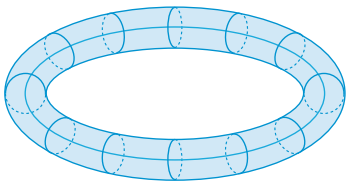
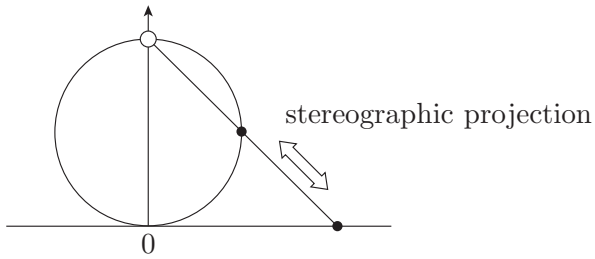


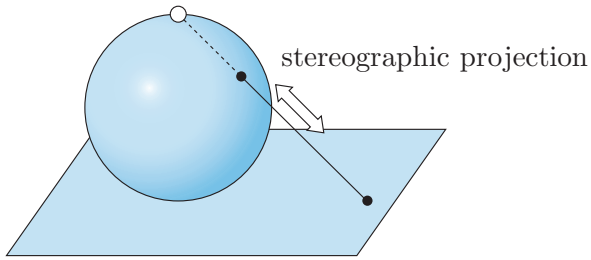
**Figure 4.3**

Now bend this around and imagine joining the two ends together seamlessly. This would produce a circular pipe (in the sense that placing the circular pipe onto sand would leave a circular imprint). The shape of this circular pipe would be similar to that of a child's rubber bathing ring (Figure 4.3). If you cut through this circular pipe at any point then you would see the cross-section of the original hose – in other words, a circle. So we can think of this surface as being the product of two circles. Mathematically this surface is a torus. Therefore we can now deduce that the torus is connected.

The punctured circle is homeomorphic to the Euclidean line $\mathbb{R}$, by means of stereographic projection (Figure 4.4(a)). Similarly, the punctured sphere is homeomorphic to the Euclidean plane (Figure 4.4(b)). Since the Euclidean plane is connected, so is the punctured sphere. The puncture is repaired by the operation of *closure*, and since closure preserves connection, by Theorem 3.12, the sphere is connected.




(a)




(b)

**Figure 4.4**

# 5 Path-connected spaces

Basing our notion of connectedness on whether or not a space falls into pieces seems very natural if we consider looking at a space from the 'outside'. However, if we lived *inside* the space, then we might come up with an alternative concept. We may decide that a space is connected if we can get from any one point to any other point without leaving our space. This alternative description of being connected leads to the idea of *path-connectedness*. Intuitively we say that a space is *path-connected* if it is possible to find a 'path' between any two of its points (Figure 5.1).

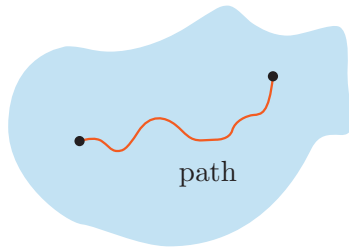



**Figure 5.1**

We cannot proceed any further until we define what we mean by a *path* in a metric space. The concept of a path is of wide interest. If you have studied a course in complex analysis (contour integration), fluid mechanics (vortices), electromagnetism (the integral laws relating currents and magnetic fields), or differential geometry and its applications to general relativity, you have met path (or line) integrals, and hence paths. Think for a moment about how you could define a path ....

... You may have thought about something that you can draw without taking your pen off the page. This is a useful way of thinking about a path in $\mathbb{R}^2$, but not so helpful when considering a path in, say, a space of functions. Writing definitions is often tricky and mathematicians spend a fair amount of their time trying to correctly define their terms. Indeed, getting the definitions right can be the hardest part of proving a result.

## 5.1 Defining a path

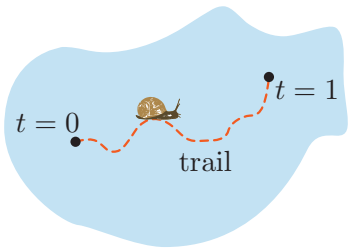



**Figure 5.2**

What should we mean by a path? We certainly want our idea of a mathematical path to include the intuitive idea of a path on the plane, so we start there. Imagine a snail crawling around in $\mathbb{R}^2$ (Figure 5.2). As the snail moves, it leaves a trail behind it on the surface. Suppose the motion takes place from time $t = 0$ to $t = 1$, after which we take the snail off the surface and examine the trail left behind.

What will not be apparent is how the trail was made: did the snail speed up, or slow down? Did it stop for a rest at any point? Did it reverse back along a portion of the trail, then reverse again and go on its way? There are infinitely many ways that a given trail can be laid down, so the concept of a trail is inadequate for mathematical purposes.

A more informative way to describe the snail's progress is to record where it is at each instant in time – that is, instead of focusing on the trail, we consider the *function* that records the snail's position as a function of time. Mathematically, this is a function $p\colon [0,1] \to X$, where $X$ is the surface upon which the snail moves. The interpretation is that the snail is at the point $p(t) \in X$ at time $t \in [0,1]$: in particular, it starts at $p(0)$ and finishes at $p(1)$. In this language, the trail is the image set $p([0,1]) \subseteq X$.

The only condition we demand of a path is that it be a $(d^{(1)}_{[0,1]}, d)$-*continuous function*: the snail may speed up, slow down, stop, reverse, change direction, as often and as abruptly as it pleases, but it must not 'break' its motion by, for example, jumping along at any point.

The snail is not a grasshopper!

The Peano curve is such an example (see Figure 5.3).

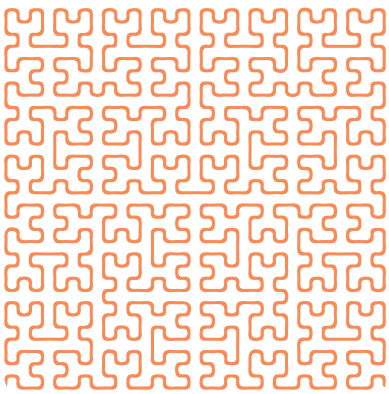


**Figure 5.3**

The trail may be extremely complicated and hard to visualise. It is known, for example, that there are paths on the plane whose trails completely fill the unit square! However, the advantage of thinking of a path as a function is that the fact that our path must not have any breaks in it can be encapsulated in the idea that the function must be continuous. If we focused on the image set (the trail) then we could possibly come up with a workable definition for the plane, but it would be harder to generalise our definition to paths in other spaces. With this definition, there is no reason as to why we should confine ourselves to paths on surfaces in $\mathbb{R}^2$. All that we require is continuity, and so everything we have said so far holds if the codomain is an arbitrary metric space (except, of course, for the bits about the snail).

**Definition 5.1** *Path*

Let $(X, d)$ be a metric space and let $[0,1]$ have the induced Euclidean metric $d^{(1)}_{[0,1]}$.

A function $p\colon [0,1] \to X$ is a **path** in $X$ if $p$ is a $(d^{(1)}_{[0,1]}, d)$-continuous function (Figure 5.4). A path $p$ **goes from** $a$ **to** $b$ in $X$ if $p$ is a path in $X$ and if there are $t_x$, $t_y \in [0,1]$ with $t_x \leq t_y$, $p(t_x) = a$ and $p(t_y) = b$. If the order in which the points are reached is unimportant, then we say that $p$ **joins** $a$ and $b$.

The **initial point** of a path $p$ is the point $p(0)$ in $X$ and the **final point** of $p$ is the point $p(1)$.

The path is a **closed path** if the initial and final points of a path are the same (that is, $p(0) = p(1)$).

The **trail** of a path is its image set $p([0,1])$.

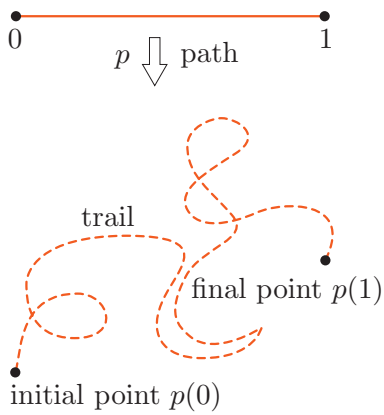



**Figure 5.4**

### Remarks

1. The closed nature of the domain ensures that our path has a definite starting point and ending point.
2. Starting with a connected domain enables us to use Lemma 3.6 to guarantee that the trail $p([0,1])$ is a connected subset of $X$.
3. The meaning of 'closed' in 'closed path' is different from its meaning in 'closed set'. This unfortunate coincidence of terminology is standard. It is usually clear from the context which meaning of 'closed' is intended.
4. For each point $x$ in a metric space $(X, d)$, there is always a path joining a point $x$ to itself – namely, the constant function $p\colon [0,1] \to X$ given by $p(t) = x$.

### Exercise 5.1

Let $(X, d)$ be a metric space. Prove part 4 of the remarks above. In other words, prove that the constant function $p\colon [0,1] \to X$ given by $p(t) = x$ is $(d^{(1)}_{[0,1]}, d)$-continuous.

### Worked Exercise 5.2

Let $\mathbb{R}^2$ have the Euclidean metric. Find a path with initial point $(0,0)$ and final point $(1,1)$.

#### Solution

The simplest example of such a path is probably the function $p\colon [0,1] \to \mathbb{R}^2$ given by $p(t) = (t,t)$, illustrated in Figure 5.5. We need to verify that this $p$ is $(d^{(1)}_{[0,1]}, d^{(2)})$-continuous. In order to do so, we note that $p_1 \circ p$ and $p_2 \circ p$ are both $(d^{(1)}_{[0,1]}, d^{(1)})$-continuous functions from $[0,1]$ to $\mathbb{R}$, where $p_1$ and $p_2$ denote the usual projection functions from $\mathbb{R}^2$ to $\mathbb{R}$. Hence, by Theorem 3.12 of Chapter 15, the function $p$ is $(d^{(1)}_{[0,1]}, d^{(2)})$-continuous. Since $p(0) = (0,0)$ and $p(1) = (1,1)$, the map $p$ gives a path with initial point $(0,0)$ and final point $(1,1)$.

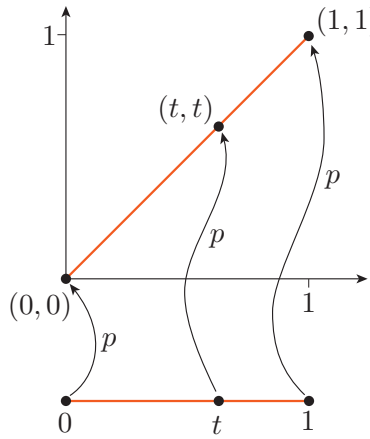



**Figure 5.5**

### Exercise 5.2

Let $X = [0,1] \times [0,1]$ with the induced Euclidean metric. Show that there is a path joining $(0,0)$ and $(1, \frac{1}{2})$ in $X$.

We next show that paths can be joined. This may seem obvious, but we need to ensure that we can join two paths together and still end up with a continuous map from $[0,1]$ to $X$, since our definition of path requires this. We first need a lemma.

### Lemma 5.3

Suppose we have maps $f\colon [a,b] \to \mathbb{R}$ and $g\colon [b,c] \to \mathbb{R}$, with $f(b) = g(b)$ and where $f$ and $g$ are continuous with respect to the metrics induced by the Euclidean metric on $\mathbb{R}$.
Then $h\colon [a,c] \to \mathbb{R}$ defined by:

$$h(t) = \begin{cases} f(t), & \text{for } a \leq t \leq b, \\ g(t), & \text{for } b \leq t \leq c, \end{cases}$$

is also continuous with respect to the metrics induced by the Euclidean metric on $\mathbb{R}$.

### Proof

Throughout this proof, we work with respect to the metrics induced by the Euclidean metric on $\mathbb{R}$. We write $d$ for $d^{(1)}$. By Theorem 7.1 of Chapter 16, the function $h\colon [a,c] \to \mathbb{R}$ will be $(d_{[a,c]}, d)$-continuous if, and only if, $h^{-1}(F)$ is a $d_{[a,c]}$-closed set in $[a,c]$ whenever $F$ is a $d$-closed set in $\mathbb{R}$.

First we note the following: a set $A \subseteq [r,s]$ is a $d_{[r,s]}$-closed subset of $[r,s]$ if, and only if, we can find a $d$-closed subset of $\mathbb{R}$, say $E$, with $A = [r,s] \cap E$. Now $[r,s]$ is $d$-closed in $\mathbb{R}$ and so $A$, being the intersection of two $d$-closed sets, is $d$-closed. So $A$ is both a $d$-closed subset of $\mathbb{R}$ and a $d_{[r,s]}$-closed subset of $[r,s]$.

Let $F$ be a $d$-closed set in $\mathbb{R}$. Then since $f(b) = g(b)$ we have:

$$\begin{aligned} h^{-1}(F) &= \{x \in [a,c] : h(x) \in F\} \\ &= \{x \in [a,b] : f(x) \in F\} \cup \{x \in [b,c] : g(x) \in F\} \\ &= f^{-1}(F) \cup g^{-1}(F). \end{aligned}$$

But $f$ is $(d_{[a,b]}, d)$-continuous and $g$ is $(d_{[b,c]}, d)$-continuous, and so by Theorem 7.1 of Chapter 16, $f^{-1}(F)$ is a $d_{[a,b]}$-closed subset of $[a,b]$ and $g^{-1}(F)$ is a $d_{[b,c]}$-closed subset of $[b,c]$. Therefore, by the note near the start of this proof, $f^{-1}(F)$ and $g^{-1}(F)$ are both $d$-closed subsets of $\mathbb{R}$ and so their union $f^{-1}(F) \cup g^{-1}(F)$ is a $d$-closed subset of $\mathbb{R}$. Hence, by Theorem 3.12 of Chapter 16, $f^{-1}(F) \cup g^{-1}(F) = f^{-1}(F) \cup g^{-1}(F) \cap [a,c]$ is a $d_{[a,c]}$-closed subset of $[a,c]$, and $h$ is $(d_{[a,c]}, d)$-continuous. ■

**Lemma 5.4**

Let $(X, d)$ be a metric space and let $a, b, c \in X$. Suppose that there is a path $p$ in $X$ with initial point $a$ and final point $b$, and a path $q$ in $X$ with initial point $b$ and final point $c$. Then there is a path that joins $a$ and $c$ in $X$; see Figure 5.6.

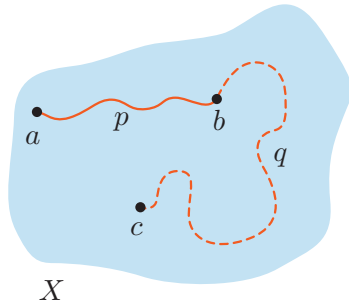



**Figure 5.6**

**Proof** In this proof, all metrics are the metrics induced by the Euclidean metric on $\mathbb{R}$.

The basic idea here is simple: we define a path joining $a$ and $c$ by doing $p$ followed by $q$. There is a problem, however: both $p$ and $q$ are defined for $t \in [0, 1]$, and what we would like is to trace the whole of $p$ for $t \in [0, \frac{1}{2}]$ and then trace the whole of $q$ for $t \in [\frac{1}{2}, 1]$, as illustrated in Figure 5.7. The way to do this is to trace out both $p$ and $q$ at double their original 'speed'. Formally, we do this as follows.

Define $r\colon [0,1] \to X$ by

$$r(t) = \begin{cases} p(2t), & \text{for } 0 \leq t \leq \frac{1}{2}, \\ q(2t-1), & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

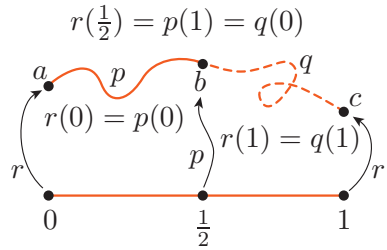



**Figure 5.7**

Certainly $r(0) = p(0) = a$ and $r(1) = q(1) = c$ and so it remains only to show that $r$ is $(d^{(1)}_{[0,1]}, d)$-continuous. The map taking $t$ to $p(2t)$ is the composition of $p$ with the continuous map $f\colon t \mapsto 2t$, and the map taking $t$ to $q(2t-1)$ is the composition of $q$ with the continuous map $g\colon t \mapsto 2t-1$. Therefore both are continuous (by the Composition Rule for continuous functions in Book D, Chapter 13). Also $r(\frac{1}{2}) = p(1) = q(0) = b$. Therefore $r$ is continuous by Lemma 5.3. ■

Now that we have the notion of a path, we can define what we mean for a metric space to be path-connected.

**Definition 5.5** *Path-connected*

Let $(X, d)$ be a metric space.

The space $(X, d)$ is **path-connected** if, for each $a$ and $b$ in $X$, there is a path in $X$ that joins $a$ and $b$.

A set $A \subseteq X$ is **path-connected** if $(A, d_A)$ is path-connected.

Recall that the induced metric on $A$ is denoted by $d_A$.

### Remarks

1. If $X = \emptyset$, then $X$ is certainly path-connected, since there are no points to join.
2. If $X = \{x\}$ has only one point, the constant function $f\colon [0,1] \to X$ given by $f(t) = x$ is a path, and so $X$ is path-connected.
3. When we wish to emphasise the metric, we use the term *d-path-connected*.

In the next exercise, we ask you to show that paths can be traced backwards.

### Exercise 5.3

Let $(X, d)$ be a metric space and let $a, b \in X$. Suppose that $p\colon [0,1] \to X$ is a path in $X$ with initial point $a$ and final point $b$. Show that there is a path $q$ in $X$ with initial point $b$ and final point $a$.

## 5.2 Examples of path-connected spaces

We now look at some examples.

### Worked Exercise 5.6

Show that $\mathbb{R}$ with the Euclidean metric is path-connected.

#### Solution

In this solution, all metrics are the metrics induced by the Euclidean metric on $\mathbb{R}$.

Let $a$, $b \in \mathbb{R}$; we must find a path joining $a$ and $b$ in $\mathbb{R}$. We already know how to define such a path if $a = b$, so suppose that $a \neq b$. Without loss of generality, we shall assume that $a < b$. There are many paths that we can choose. The standard one is the straight-line path $P\colon [0,1] \to \mathbb{R}$ defined by

$$P(t) = (1-t)a + tb,$$

illustrated in Figure 5.8. As $P$ is simply the restriction of a linear function to the interval it is clearly continuous. Also we have $P(0) = a$ and $P(1) = b$. Hence $P$ is a path in $\mathbb{R}$ from $a$ to $b$ and $\mathbb{R}$ is path-connected.

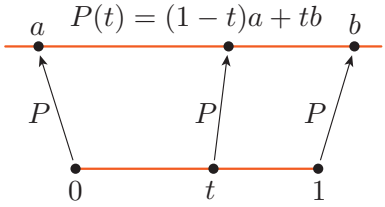



**Figure 5.8**

### Exercise 5.4

Show that $[a, b]$ with the induced Euclidean metric is path-connected.

#### Remark

The result of Exercise 5.4 enables us to deduce that every interval $I$ in $\mathbb{R}$ is path-connected. We know that the empty interval is path-connected. If $I$ is not empty, for each $a$ and $b$ in $I$ $(a \leq b)$ we know that $[a, b]$ is path-connected and hence there is a path from $a$ to $b$ in $[a, b] \subseteq I$. Thus $I$ is path-connected.

The result in Worked Exercise 5.6 generalises to $\mathbb{R}^n$ (for any $n \in \mathbb{N}$).

**Proposition 5.7**

The metric space $\mathbb{R}^n$ with the Euclidean metric is path-connected.

**Proof** For $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ we define $P\colon [0,1] \to \mathbb{R}^n$ by $P(t) = (1-t)\mathbf{a} + t\mathbf{b}$.

We must prove that $P$ is a path from $\mathbf{a}$ to $\mathbf{b}$.

*Continuity:* let $p_j\colon \mathbb{R}^n \to \mathbb{R}$ be the $j$th projection function, defined by $p_j(\mathbf{a}) = a_j$. Then $(p_j \circ P)(t) = (1-t)a_j + tb_j$, so $p_j \circ P$ is linear and so clearly continuous. Since $p_j \circ P$ is continuous for $j = 1, 2, \ldots, n$, it follows from the generalisation of Theorem 3.12 of Chapter 15 that $P$ is continuous.

*Endpoints:* by inspection, $P(0) = \mathbf{a}$ and $P(1) = \mathbf{b}$. Thus $P$ is a path from $\mathbf{a}$ to $\mathbf{b}$ and $\mathbb{R}^n$ is path-connected. ■

**Exercise 5.5**

In the proof of Proposition 5.7 we have suppressed the metrics and written 'continuous' rather than $(d, e)$-continuous (for appropriate metrics $d$ and $e$). Insert the appropriate metrics for each use of the word 'continuous' in this proof.

We have seen that connectedness is preserved under continuous mappings, and hence is a topological invariant. The same is true for path-connectedness.

**Theorem 5.8**

Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces, let $X$ be path-connected, and let $f\colon X \to Y$ be $(d_X, d_Y)$-continuous. Then $f(X)$ is path-connected.

**Proof** Let $a, b \in f(X)$. By the definition of $f(X)$, there exist $c, d \in X$ such that $f(c) = a$ and $f(d) = b$. Since $X$ is path-connected, there exists a path $p\colon [0,1] \to X$ from $c$ to $d$. Since $f$ is $(d_X, d_Y)$-continuous and $p$ is $(d^{(1)}_{[0,1]}, d_X)$-continuous, $f \circ p\colon [0,1] \to f(X)$ is a path from $f(c) = a$ to $f(d) = b$ in $f(X)$ (with the induced metric inherited from $(Y, d_Y)$). Thus $f(X)$ is path-connected. ■

**Corollary 5.9**

Path-connectedness is a topological invariant.

We can use the methods of the proof of Proposition 5.7 to show that every $d^{(n)}$-open ball in $\mathbb{R}^n$ is path-connected for the Euclidean metric. The first step is to show that the unit open ball centred at $\mathbf{0}$ in $\mathbb{R}^n$ is path-connected for the Euclidean metric. We then use the fact that path-connectedness is a topological invariant to show that this applies to every $d^{(n)}$-open ball.

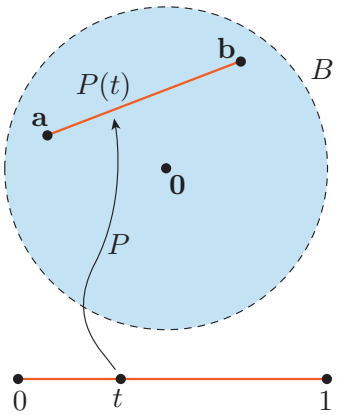



**Figure 5.9**

### Worked Exercise 5.10

Show that the unit $d^{(n)}$-open ball $B$ in $\mathbb{R}^n$ with the induced Euclidean metric is path-connected.

#### Solution

Let $\mathbf{a}, \mathbf{b} \in B$. We know from the proof of Proposition 5.7 that $P(t) = (1-t)\mathbf{a} + t\mathbf{b}$ defines a path from $\mathbf{a}$ to $\mathbf{b}$ in $\mathbb{R}^n$ (Figure 5.9). Therefore we can deduce that $P$ defines a path in $B$ if we can show that $P(t)$ belongs to $B$ for all $t \in [0,1]$ – that is, we need to show that $d^{(n)}\big(P(t), \mathbf{0}\big) < 1$. We have:

$$\begin{aligned} d^{(n)}\big(P(t), \mathbf{0}\big) &= d^{(n)}\big((1-t)\mathbf{a} + t\mathbf{b}, \mathbf{0}\big) \quad \text{(definition of } P\text{)} \\ &\leq d^{(n)}\big((1-t)\mathbf{a} + t\mathbf{b}, t\mathbf{b}\big) + d^{(n)}\big(t\mathbf{b}, \mathbf{0}\big) \quad \text{(Triangle Inequality)} \\ &= d^{(n)}\big((1-t)\mathbf{a}, \mathbf{0}\big) + d^{(n)}\big(t\mathbf{b}, \mathbf{0}\big) \quad \text{(translating by } t\mathbf{b}\text{)} \\ &= (1-t)d^{(n)}\big(\mathbf{a}, \mathbf{0}\big) + td^{(n)}\big(\mathbf{b}, \mathbf{0}\big) \quad \text{(factorising)} \\ &< (1-t) + t = 1. \quad (\mathbf{a}, \mathbf{b} \in B) \end{aligned}$$

Therefore $B$ is path-connected.

A proof almost identical to that in Worked Exercise 5.10 allows us to show that the unit $d^{(n)}$-closed ball in $\mathbb{R}^n$ is path-connected, and thus to deduce that every $d^{(n)}$-closed ball in $\mathbb{R}^n$ is path-connected.

### Theorem 5.11

Every $d^{(n)}$-open ball in $\mathbb{R}^n$ is path-connected. Every $d^{(n)}$-closed ball in $\mathbb{R}^n$ is path-connected.

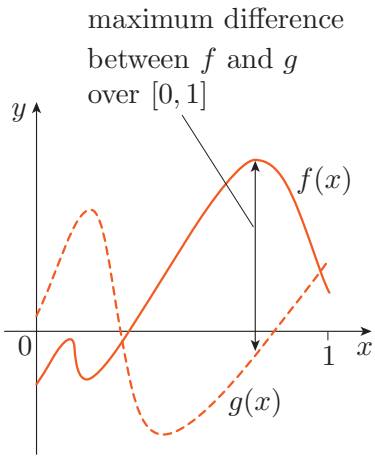



**Figure 5.10**

So far, we have just considered path-connectedness for the Euclidean metric. We now consider a different metric space.

Recall that $C[0,1]$ is the set of continuous functions from $[0,1]$ to $\mathbb{R}$ and that the max metric on $C[0,1]$ is defined by

$$d_{\max}(f,g) = \max\{|g(x) - f(x)| : x \in [0,1]\}.$$

It may seem a little strange to think of a path in a space of functions but there is no mathematical reason not to. Our metric, $d_{\max}$, gives us a sense of the distance between two functions, where the distance between $f$ and $g$ is the maximum vertical distance between them (as shown in Figure 5.10).

One function is 'close' to another if the maximum vertical distance between the two graphs of the functions is 'small'. Intuitively, we can think of a trail, $p\big([0,1]\big)$, from the function $f$ to the function $g$ as being a series of functions, $p(t) \in C[0,1]$, such that drawing the graphs of each would give us a way of deforming the graph of $f$ into the graph of $g$. This is not a rigorous definition, but may be a helpful way of visualising a trail in a function space.

### Worked Exercise 5.12

Show that $\big(C[0,1], d_{\max}\big)$ is path-connected.

#### Solution

Let $f$ and $g$ be functions in $C[0,1]$. Consider the function $P\colon [0,1] \to C[0,1]$ given by

$$P(t) = (1-t)f + tg,$$

so that, for $x \in [0,1]$ and $t \in [0,1]$,

$$P(t)(x) = (1-t)f(x) + tg(x).$$

We ask you to show that $P$ is $(d^{(1)}_{[0,1]}, d_{\max})$-continuous in Exercise 5.6. Note that $P(0) = f$ and $P(1) = g$. We deduce that $P$ is a path from $f$ to $g$, and $C[0,1]$ is path-connected.

### Exercise 5.6

By bounding $d_{\max}\big(P(t), P(s)\big)$ in terms of $|t-s|$ and $d_{\max}(f,g)$, prove that $P$ is $(d^{(1)}_{[0,1]}, d_{\max})$-continuous.

You may have noticed that all the examples of path-connected spaces that you have met so far are also connected. This is no accident, as the following theorem shows.

### Theorem 5.13

Let $(X,d)$ be a metric space. If $X$ is $d$-path-connected, then $X$ is $d$-connected.

**Proof** We need to show that if we can find a path between any two points in our space then we cannot form a disconnection. If $X$ is empty, then there is nothing to prove, so we can suppose that $X \neq \emptyset$.

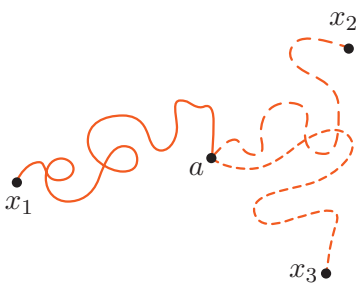



**Figure 5.11**

Fix some $a \in X$. If $X$ is path-connected, then each point $x \in X$ can be joined to $a$ by some path in $X$ (Figure 5.11). Each such path defines a trail. The trail is the continuous image of the connected set $[0, 1]$ and so, by Lemma 3.6 the trail is a connected set, containing both $a$ and $x$. The union of all these connected trails is the whole of $X$ (as every point in $X$ is the endpoint of at least one trail), and the intersection is non-empty since each trail contains the point $a$. It follows from Lemma 3.15 that the union is a connected set and so $X$ is connected. ■

### Corollary 5.14

The space $\big(C[0, 1], d_{\max}\big)$ is connected.

**Proof** This follows from Theorem 5.13 and Worked Exercise 5.12. ■

It follows from Theorem 5.13 that one way to show that a space is connected is to show that it is path-connected.

### Worked Exercise 5.15

Let $A \subseteq \mathbb{R}^2$ be a countable set. Show that, for the induced Euclidean metric, $\mathbb{R}^2 - A$ is path-connected, and hence connected.

#### Solution

Since $\mathbb{R}^2$ contains a copy of $\mathbb{R}$, we know that $\mathbb{R}^2$ is uncountable. Also $A$ is countable. Therefore we have $\mathbb{R}^2 - A$ uncountable (since $\mathbb{R}^2$ is the union of $\mathbb{R}^2 - A$ and $A$, and the union of two countable sets is countable by Book D, Chapter 16, Corollary 6.10).

Let $\mathbf{a}$ and $\mathbf{b}$ be distinct points of $\mathbb{R}^2 - A$. We describe how to find a path joining $\mathbf{a}$ and $\mathbf{b}$. Let $L$ be the perpendicular bisector of the line segment joining $\mathbf{a}$ and $\mathbf{b}$. $L$ is uncountable since it is a copy of $\mathbb{R}$. For each $\mathbf{x} \in L$, let $p_{\mathbf{x}} \colon [0, 1] \to \mathbb{R}^2$ be a path joining $\mathbf{a}$ to $\mathbf{b}$ with trail as illustrated in Figure 5.12. Let $\mathbf{x}$ and $\mathbf{y}$ be distinct points in $L$; then the trails of their corresponding paths $p_{\mathbf{x}}\big([0, 1]\big)$ and $p_{\mathbf{y}}\big([0, 1]\big)$ meet only at the points $\mathbf{a}$ and $\mathbf{b}$. In particular, $A \cap p_{\mathbf{x}}\big([0, 1]\big)$ is disjoint from $A \cap p_{\mathbf{y}}\big([0, 1]\big)$ whenever $\mathbf{x} \neq \mathbf{y}$. Since $L$ is uncountable and $A$ is countable, this means that there is $\mathbf{x} \in L$ for which $p_{\mathbf{x}}\big([0, 1]\big) \cap A = \emptyset$. For this $\mathbf{x}$, the map $p_{\mathbf{x}}$ is a path in $\mathbb{R}^2 - A$ joining $\mathbf{a}$ to $\mathbf{b}$. Thus $\mathbb{R}^2 - A$ is path-connected. It follows from Theorem 5.13 that $\mathbb{R}^2 - A$ is connected.

Countable sets were introduced in Book D, Chapter 16.

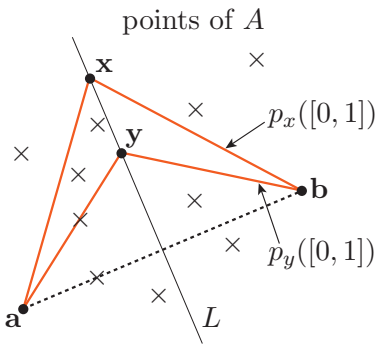



**Figure 5.12**

### Remarks

1. This argument works for any $\mathbb{R}^n$ with $n > 1$ but fails for $\mathbb{R}$ (as we do not have enough room to construct $L$ in one-dimensional space). In fact, as we know from Exercise 4.5, removing just one point from $\mathbb{R}$ gives a disconnected set!

2. Since the empty set is countable, this result confirms that $\mathbb{R}^2$ is path-connected.

### Exercise 5.7

Show that $\mathbb{R}^2$ is not homeomorphic to $\mathbb{R}$.

### Exercise 5.8

Let $X$ be a set containing at least two points and let $d_0$ be the discrete metric on $X$. Show that $X$ is not path-connected.

(*Hint*: use the result of Exercise 3.4.)

## 5.3 Products of path-connected spaces

We know that taking products preserves connectedness, so it is natural to ask whether taking products preserves path-connectedness. The answer is yes.

### Theorem 5.16

Let $(X, d_X)$ and $(Y, d_Y)$ be non-empty metric spaces. The product space $(X \times Y, e)$, where $e$ is a product metric, is path-connected if, and only if, $(X, d_X)$ and $(Y, d_Y)$ are both path-connected.

The proof is similar to that given for Theorem 4.8, and we omit it.

As with connectedness, this product result tell us that, for any intervals $I_1, \ldots, I_n$ of $\mathbb{R}$, $I_1 \times \cdots \times I_n$ is path-connected. In particular, if $I_k = \mathbb{R}$ for each $k$, we have another proof that $\mathbb{R}^n$ is path-connected.

We now turn our attention to unions and prove a result corresponding to Lemma 3.15 for connected spaces.

### Theorem 5.17

Let $(X, d)$ be a metric space, and let $\{A_i : i \in I\}$ be a family of path-connected subsets of $X$ whose intersection is non-empty. Then $A = \bigcup_{i \in I} A_i$ is path-connected.

**Proof** Let $x, y \in A$; we want to find a path in $A$ joining $x$ and $y$.

Since $\bigcap_{i \in I} A_i \neq \emptyset$, we can fix a point $a \in \bigcap_{i \in I} A_i$ (Figure 5.13).

Since $x \in A$, there is an $i$ so that $x \in A_i$. But $a \in A_i$ and so there is a path connecting $x$ to $a$. Similarly, we can find a path joining $y$ to $a$. We now use the result of Exercise 5.3 together with Lemma 5.4 to find a path joining $x$ to $y$, proving that $A$ is path-connected. ■

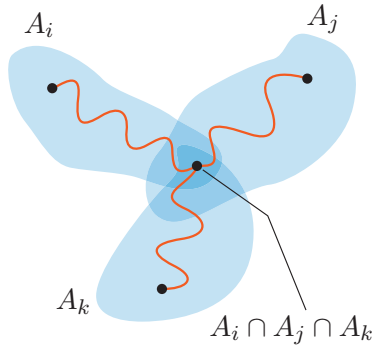



**Figure 5.13**

You may be wondering why we are so interested in path-connectedness when it seems so similar to connectedness. This is because there are differences between the two concepts. A path-connected space is always connected but *a connected space is not always path-connected*. In other words, just because a point inhabits the same connected space as another does not mean that it is possible to get from one to the other if you live in that space. In the next section we examine a space that is connected but not path-connected. In addition we will see that, unlike connectedness, path-connectedness is not preserved under closure.

# 6 The topologist's cosine

It seems reasonable to conjecture that connectedness implies path-connectedness, and that path-connectedness is preserved by closure. However, it is not possible to prove these conjectures because they are false! When it turns out to be difficult to *prove* a conjecture, then the next step is to look for a counter-example (and so *disprove* the conjecture). We do this in this section.

Constructing counter-examples is an important part of doing mathematics.

When constructing counter-examples it is often helpful to think about what could 'go wrong'. In this case, what are the possible obstructions to finding a path between two points in connected space? A path is a function from the closed interval $[0,1]$ to a space; the interval $[0,1]$ is not homeomorphic to $\mathbb{R}$, whereas $(0,1)$ is homeomorphic to $\mathbb{R}$. This suggests that, in some sense, $[0,1]$ cannot be 'stretched' as much as $(0,1)$. Intuitively, this is a consequence of the 'hard edges' of the closed interval: the image of $[0,1]$ under a continuous map needs to 'stop' somewhere, whereas the 'fuzzy' edges of $(0,1)$ can be smeared out indefinitely. We exploit this limitation of $[0,1]$ in our counter-example by constructing a connected space containing two points which are 'too far apart' to be reached via a path.

In the next chapter we investigate compactness and discover that $[0,1]$ is compact, whereas $(0,1)$ is not compact.

Our inspiration comes from the graph of the $\cos(\pi/x)$ function, which we investigated in Book D, Chapter 13, Worked Exercise 2.6. We defined this function $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} \cos(\pi/x), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

In fact, it is not possible to redefine this function at 0 to make it continuous (with respect to the Euclidean metric).

Recall that this function, which oscillates very rapidly close to the origin, is continuous everywhere except at 0. The graph of $f$ includes points arbitrarily close to each point on the line segment $\{0\} \times [-1,1]$ but only actually includes the point $(0,0)$ from this line segment.

If we consider the graph of the function $f$ restricted to $(0,1]$, we have the set

$$E = \{(x, \cos(\pi/x))\colon 0 < x \leq 1\} \subset \mathbb{R}^2,$$

shown in Figure 6.1. This set (the graph $E$) is path-connected for its induced Euclidean metric. Once we specify any two points, $a$ and $b$, on the

graph, we can bridge the gap between them using a continuous image of $[0, 1]$ (as there will always be finitely many oscillations between the specified points).

So we need to work a little harder to find our counter-example. We know that the graph does not include the points on the $y$-axis, so including these in our set may give us our counter-example. We also know that the graph gets arbitrarily close to the points on the $y$-axis, so it is likely that the closure of our graph, $\mathrm{Cl}_{(\mathbb{R}^2, d^{(2)})}(E)$, includes the vertical line segment $B = \{0\} \times [-1, 1]$ on the $y$-axis, shown in Figure 6.2. In fact, $\mathrm{Cl}_{(\mathbb{R}^2, d^{(2)})}(E) = E \cup B$.

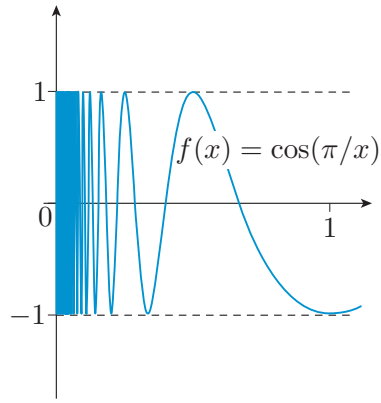



**Figure 6.1**

We know that connectedness is preserved under taking set closures so $\mathrm{Cl}_{(\mathbb{R}^2, d^{(2)})}(E)$ is connected. Is this new set path-connected? The answer is no. We obtain an idea as to why not if we try to visualise a trail between a point in $B$ and a point in $E$. Our trail would need to cover 'infinitely many' large and dense oscillations, which is not possible.

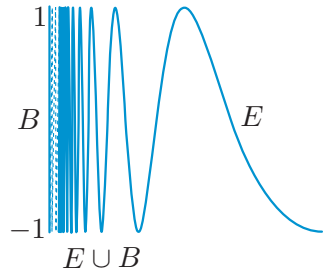



**Figure 6.2**

In order to simplify the calculations, our example, inspired by the $\cos(\pi/x)$ function, is constructed from straight line segments. This makes it much easier to determine whether a given point in the plane lies in the set.

Working out all the details in this example is a long process. Most of it is what pure mathematicians call *routine*; a word that should not be confused with *easy* or *elementary*. Routine work consists of providing details of standard pieces of analysis, such as proving the continuity of certain functions, or finding an upper bound for a set. We have concentrated on explaining the gist of the argument in the sketch proof that follows and have left some of the details as exercises. You may like to leave these exercises for now and return to them as part of your revision of Book F.

We begin by constructing a straight-line version of the set $E$ we have just described.

For $n \in \mathbb{N}$, let $A_n$ denote the set in the plane consisting of the line segment joining the point $(\frac{1}{n+1}, 0)$ to the point $(\frac{1}{2}(\frac{1}{n} + \frac{1}{n+1}), 1)$, together with the line segment joining $(\frac{1}{2}(\frac{1}{n} + \frac{1}{n+1}), 1)$ to $(\frac{1}{n}, 0)$, as shown in Figure 6.3.

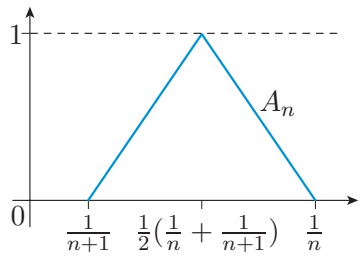



**Figure 6.3**

The algebraic description is unpleasant, but for future reference we give it:

$$A_n = \left\{(s, 2n(n+1)s - 2n) \colon \tfrac{1}{n+1} \leq s \leq \tfrac{1}{2}\left(\tfrac{1}{n} + \tfrac{1}{n+1}\right)\right\} \cup \left\{(s, 2(n+1) - 2n(n+1)s) \colon \tfrac{1}{2}\left(\tfrac{1}{n} + \tfrac{1}{n+1}\right) \leq s \leq \tfrac{1}{n}\right\}.$$

Let $A$ be the plane subset

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Then $A$ is the *graph* of the function $l \colon (0, 1] \to \mathbb{R}$ given by

$$l(x) = \begin{cases} 2n(n+1)x - 2n, & \text{if } \frac{1}{n+1} < x \leq \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right), \ n \in \mathbb{N}, \\ 2(n+1) - 2n(n+1)x, & \text{if } \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right) < x \leq \frac{1}{n}, \ n \in \mathbb{N}, \end{cases}$$

This is similar to describing the familiar U-shaped quadratic curve as the graph of $f$, where $f(x) = x^2$, and the image set of $F$, where $F(x) = (x, x^2)$.

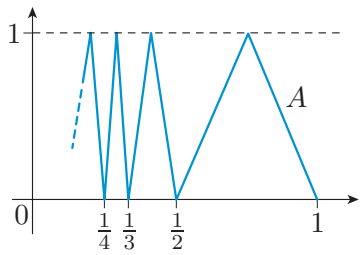



**Figure 6.4**

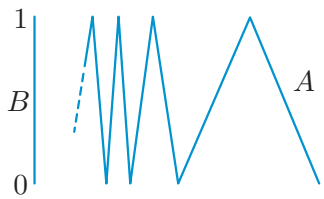



**Figure 6.5**

and shown in Figure 6.4. It is also the *image set* of the function

$$L\colon (0,1] \to \mathbb{R}^2 \text{ given by } L(x) = (x, l(x)).$$

We write $d_A$ for $d_A^{(2)}$, the metric on $A$ induced by the Euclidean metric on $\mathbb{R}^2$.

Let $B$ be the vertical line segment joining the origin to the point $(0,1)$:

$$B = \{(0,y)\colon 0 \leq y \leq 1\}.$$

The *topologist's cosine* is the plane set

$$C = A \cup B,$$

consisting of all the points of $A$ and $B$ (Figure 6.5). We write $d_C$, for the metric on $C$ induced by the Euclidean metric for the plane.

### Lemma 6.1

The function $l\colon (0,1] \to \mathbb{R}$ is $(d_{(0,1]}^{(1)}, d^{(1)})$-continuous and the function $L\colon (0,1] \to \mathbb{R}^2$ is $(d_{(0,1]}^{(1)}, d^{(2)})$-continuous.

**Proof** We ask you to prove this in Exercise 6.1. ■

This allows us to show that $A$ is a path-connected, and hence connected, subset of the plane, as follows.

### Corollary 6.2

$A$ is path-connected for $d_A$, the metric induced by the Euclidean metric on the plane.

**Proof** The set $A$ is the image set of the path-connected set $(0,1]$ under the $(d_{(0,1]}^{(1)}, d^{(2)})$-continuous function $L$. So, by Theorem 5.8, $A$ is path-connected. ■

We now know that the zigzag bit, $A$, of our set $C = A \cup B$ is $d_A$-connected. We also know that connectedness is preserved by closures (from Theorem 3.12). Therefore we can show that $C$ is a $d_C$-connected subset of the plane by showing that $C = \mathrm{Cl}(A)$. To simplify what follows, we work in $C$ and show that $C = \mathrm{Cl}_{(C,d_C)}(A)$ as this is sufficient to show that $C$ is connected (by Theorem 3.12).

### Theorem 6.3

$C$ is $d_C$-connected, where $d_C$ is the metric induced by the Euclidean metric on the plane.

**Proof** In order to show that $C$ is $d_C$-connected, we show that the closure of $A$, when considered as a subset of $C$, is $C$.

Since the whole space in question is $C$, it is certainly the case that $\mathrm{Cl}_{(C,d_C)}(A) \subseteq C$. From Theorem 3.8 in Chapter 16 we know that $\mathrm{Cl}_{(C,d_C)}(A) = \{\mathbf{a} \in C : \mathbf{a} \text{ is a } d_C\text{-closure point of } A \text{ in } C\}$. So it remains only to show that each point $\mathbf{b}$ in $C - A = B$ is a $d_C$-closure point of $A$ in $C$.

By Book D, Chapter 16, Worked Exercise 3.7, it is sufficient to find a sequence of points $(\mathbf{a}_n)$, with $\mathbf{a}_n \in A$ for each $n \in \mathbb{N}$ that is $d_C$-convergent to $\mathbf{b}$.

The point $\mathbf{b}$ is a point in $B$, and so can be written as $(0, b)$ for some $b$ with $0 \leq b \leq 1$. We find a sequence of points, $(\mathbf{a}_n)$ in $A$, with limit $\mathbf{b}$, as demonstrated in Figure 6.6. In Exercise 6.2 we ask you to prove that $\mathbf{a}_n$ is $d_C$-convergent to $\mathbf{b}$ as $n \to \infty$. Hence $\mathbf{b} \in \mathrm{Cl}_{(C,d_C)}(A)$. Thus, since $\mathbf{b}$ was an arbitrary point in $C - A$, we have that $\mathrm{Cl}_{(C,d_C)}(A) = C$, and so $C$ is $d_C$-connected by Theorem 3.12. ■

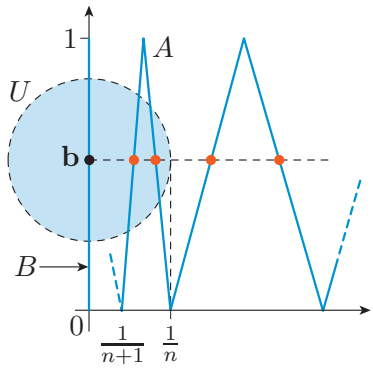



**Figure 6.6**

Thus we have a connected set, $C$, which is the closure of a path-connected set. It remains to show that $C$ is not path-connected.

**Theorem 6.4**

$C$ is not path-connected for the Euclidean metric.

**Proof** We have already observed (Corollary 6.2) that $A$ is path-connected. Since $B$ is also path-connected, the difficulty must lie in finding a path from a point in $B$ to a point in $A$. The proof is by contradiction. All metrics are those induced by the relevant Euclidean metric on $\mathbb{R}^n$.

Suppose that $C$ is path-connected; in particular, assume that there is a continuous function $P : [0, 1] \to C$ with initial point $P(0) = (0, 0)$ in $B$ and final point $P(1) = (1, 0)$ in $A$, as shown in Figure 6.7. To describe the function $P$, we use the language of the path of a particle as a function of time. At time $t = 0$, the particle is at the point $(0, 0)$, at $t = 1$ it is at $(1, 0)$, and at time $t$ it is at the point $P(t)$, somewhere in $\mathbb{R}^2$.

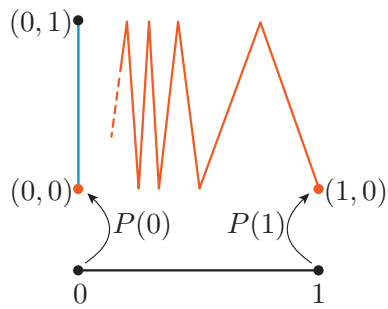



**Figure 6.7**

We used this language, with a snail instead of a particle, at the start of Section 4.

Write $P(t) = (P_1(t), P_2(t))$.

This introduces two new continuous functions, $P_1 : [0, 1] \to [0, 1]$ and $P_2 : [0, 1] \to [0, 1]$. In our particle–time language, $P_1(t)$ is the $x$-coordinate of the particle at time $t$, and $P_2(t)$ is its $y$-coordinate.

Now comes the key to the proof, certainly not an obvious idea. Since the particle starts at a point in the set $B$ and ends up at a point in the set $A$, then there must be a *latest time* $\tau$, *after which it never returns to* $B$. We find this time as follows.

If the particle kept popping back to $B$ then it would not eventually come to rest in $A$ at time $t = 1$.

Whenever the particle is in $B$, its $x$-coordinate is 0, so it is in $B$ for those times $t$ for which $P_1(t) = 0$, and for no other times. We collect all of these times together in the set $P_1^{-1}(\{0\}) \subseteq [0,1]$.

Note how this example depends crucially on the least upper bound property of $\mathbb{R}$.

The set $\{0\}$ is a $d^{(1)}$-closed subset of the $x$-axis and $P_1$ is continuous, so $P_1^{-1}(\{0\})$ is a $d^{(1)}$-closed subset of $[0,1]$. Since it is a subset of $[0,1]$ it is a bounded subset of $\mathbb{R}$, and so $P_1^{-1}(\{0\})$ has a least upper bound; call this number $\tau$.

Note that $\tau < 1$, since at $t = 1$ the particle is known to be at the point $(1,0)$ in $A$ and not in $B$.

Since $P_1^{-1}(\{0\})$ is $d^{(1)}$-closed, the supremum is contained in the set and so $P_1(\tau) = 0$. Therefore $P(\tau) \in B$ (as $\big(P_1(\tau), P_2(\tau)\big) = (0, P_2(\tau))$) and, for all later times $t$ such that $1 \geq t > \tau$, we have $P(t) \in A$ (by our choice of $\tau$).

*To summarise:* we have assumed that we have a path, $P(t) = \big(P_1(t), P_2(t)\big) \colon [0,1] \to C$, and have established that $\{P_1^{-1}\{0\}\}$ contains the supremum $\tau \in [0,1]$. We have deduced that $P_1(\tau) = 0$.

We now turn our attention to $P_2$. Remember that $P_1(t)$ is the $x$-coordinate of the path that we are trying to prove cannot exist because it oscillates so quickly. This oscillation is the key to our proof. We show that the $y$-value, $P_2(\tau)$, must simultaneously be every value in $[0,1]$ (which is, of course, impossible) and so gives a contradiction.

Since $P_2$ is continuous, for each $a \in [0,1]$, the set $P_2^{-1}(\{a\})$ is a $d^{(1)}$-closed subset of $[0,1]$. Therefore if, for some $a \in [0,1]$, we show that any $d^{(1)}$-open ball $B(\tau, r)$ intersects the $d^{(1)}$-closed set $P_2^{-1}(\{a\})$, then $\tau$ would be in the $d^{(1)}$-closure $\mathrm{Cl}(P_2^{-1}(\{a\}) = P_2^{-1}(\{a\})$; and so $P_2(\tau) = a$.

So, let $a \in [0,1]$. We claim that for each $r > 0$, there is $t \in B(\tau, r)$ for which $P_2(t) = a$.

To verify this claim, we fix $a \in [0,1]$, and let $B(\tau, r)$ be an open ball around $\tau$. Then there is $\delta > 0$ such that

Note that we place no conditions on $a$ or $r$ (other than $r > 0$).

$$[\tau, \tau + \delta) \subset B(\tau, r).$$

We are looking for an element of $B(\tau, r)$ which $P_2$ maps to $a$ so it suffices to find $s \in [\tau, \tau + \delta)$ for which $P_2(s) = a$.

We noted earlier that

$$0 = P_1(\tau) < P_1(\tau + \delta) \leq 1,$$

If $P_1(\tau + \delta) = 0$, then $\tau$ is not the least upper bound of $P_1^{-1}(\{0\})$.

and so we can find $N \in \mathbb{N}$ for which

$$0 = P_1(\tau) < \frac{1}{N} < P_1(\tau + \delta) \leq 1.$$

However, $P_1$ is continuous, and so, by the Intermediate Value Theorem (Theorem 4.6), there is $t_0 \in (\tau, \tau + \delta)$ for which

$$P_1(t_0) = \frac{1}{N},$$

as illustrated in Figure 6.8. As all of these points are in $A$, the way that we defined $A$ ensures that $P_2(t_0) = 0$. Since

$$0 = P_1(\tau) < P_1(t_0) = \frac{1}{N} \quad \text{and} \quad \frac{1}{2}\left(\frac{1}{N} + \frac{1}{N+1}\right) < \frac{1}{N},$$

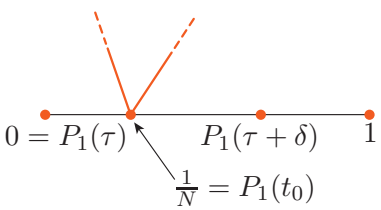



**Figure 6.8**

we can again use the Intermediate Value Theorem for $P_1$ to find $\tau < t_1 < t_0$ for which

$$P_1(t_1) = \frac{1}{2}\left(\frac{1}{N} + \frac{1}{N+1}\right),$$

and so, because of the way that $A$ is defined, $P_2(t_1) = 1$ (see Figure 6.9).

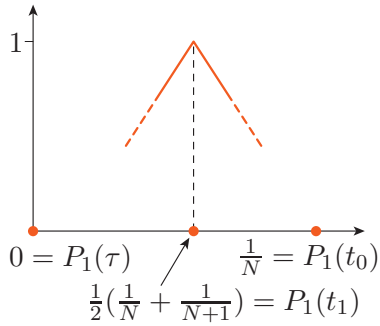



**Figure 6.9**

*To summarise:* for any $a \in [0,1]$ we want to show that any open ball around $\tau$ intersects $P_2^{-1}(\{a\})$. As $P_1(\tau) = 0$, and $P_1(\tau+\delta) > \frac{1}{N}$, we can use the Intermediate Value Theorem to find a $t_0$ with $P_1(t_0) = \frac{1}{N}$. We note that $P_2(t_0) = 0$. Here we are using the Intermediate Value Theorem to find a 'bottom corner' of the zigzag inside our $r$-ball around $\tau$. We then use the same technique to find a 'top corner' of the zigzag inside our $\varepsilon$-ball, corresponding to the point $t_1$.

Thus $\tau < t_1 < t_0 < \tau + \delta$, $P_2(t_1) = 1$ and $P_2(t_0) = 0$. But $P_2$ is continuous, and so we can now use the Intermediate Value Theorem for $P_2$ to find $t_1 \leq s \leq t_0$ with $P_2(s) = a$. So we have proved that for each $B(\tau, r)$, there is $t \in B(\tau, r)$ for which $P_2(t) = a$ and so, since $P_2^{-1}(\{a\})$ is $d^{(1)}$-closed, we must have $\tau \in P_2^{-1}(\{a\})$. Therefore $P_2(\tau) = a$. However, *we put no conditions upon* $a$, other than to require $a \in [0,1]$. This gives our required contradiction, since we cannot have $P_2(\tau) = a$ for all $a \in [0,1]$!

*To summarise:* we deduce that $P_2(t_0) = 0$ and $P_2(t_1) = 1$ for two points, $t_0$ and $t_1$, inside our $r$-ball around $\tau$. We then use the Intermediate Value Theorem again, this time applying it to $P_2$, to show that there must be a point $s$ in $B(\tau, r)$ with $P_2(s) = a$. As every open ball around $\tau$ intersects the closed set $P_2^{-1}(\{a\})$, the point $\tau$ is a closure point of this set and so must be in the set, showing that $P_2(\tau) = a$ for any $a \in [0,1]$ which is impossible.

Hence $C$ is not path-connected. ■

We have thus demonstrated that the topologist's cosine is a connected set that is not path-connected.

## Exercise 6.1

Prove Lemma 6.1, which states that the function $l\colon (0,1] \to \mathbb{R}$ is $(d^{(1)}_{(0,1]}, d^{(1)})$-continuous and the function $L\colon (0,1] \to \mathbb{R}^2$ is $(d^{(1)}_{(0,1]}, d^{(2)})$-continuous.

## Exercise 6.2

For $n \in \mathbb{N}$, let $A_n$ denote the set in the plane consisting of the line segment joining the point $(\frac{1}{n+1}, 0)$ to the point $(\frac{1}{2}(\frac{1}{n} + \frac{1}{n+1}), 1)$, together with the line segment joining $(\frac{1}{2}(\frac{1}{n} + \frac{1}{n+1}), 1)$ to $(\frac{1}{n}, 0)$, as shown in Figure 6.3.

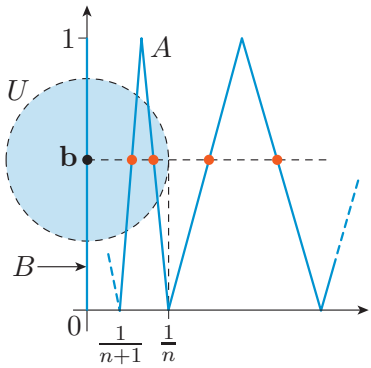



**Figure 6.10**

$$A_n = \left\{(s, 2n(n+1)s - 2n) \colon \tfrac{1}{n+1} \le s \le \tfrac{1}{2}\left(\tfrac{1}{n} + \tfrac{1}{n+1}\right)\right\}$$
$$\cup \left\{(s, 2(n+1) - 2n(n+1)s) \colon \tfrac{1}{2}\left(\tfrac{1}{n} + \tfrac{1}{n+1}\right) \le s \le \tfrac{1}{n}\right\}.$$

Let $A$ be the plane subset

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Let $B$ be the vertical line segment joining the origin to $(0, 1)$:

$$B = \{(0, y) \colon 0 \le y \le 1\}.$$

Let $C = A \cup B$ with $d_A, d_B$ and $d_C$ being the metrics on $A, b$ and $C$ respectively induced by the Euclidean metric on $\mathbb{R}$. Let $\mathbf{b} = (0, b) \in B$, so $0 \le b \le 1$. Show that we can find a sequence of points in $A$ with limit $\mathbf{b}$ as follows (see Figure 6.10). Let

$$x_n = \frac{b + 2n}{2n(n+1)}.$$

(a) Let $l$ be as defined in Exercise 6.1. Show that $l(x_n) = b$ and hence that there is at least one point in $A_n$ (namely $\mathbf{a}_n = (x_n, b)$) with $y$-coordinate equal to $b$.

(b) Let $r > 0$. Show that we can choose $n$ so that the point $\mathbf{a}_n$ is inside $B_{d_C}(\mathbf{b}, r)$.

(c) Complete the proof that $\mathrm{Cl}_{(C,d_C)}(A) = C$.

# Learning outcomes

We list those topics on which we may set assessment questions to test your understanding of the chapter.

After working through this chapter you should be able to:

(a) Define the terms *homeomorphism* and *topological invariant*.

(b) Understand the relationship between the closed and open sets in a metric space and a subspace of the metric space.

(c) Define the terms *disconnection*, *disconnected* and *connected*.

(d) Determine whether a given simple metric space is connected.

(e) Produce disconnections of some disconnected spaces.

(f) Specify which subsets of $\mathbb{R}$ are connected.

(g) Appreciate the generalisation of the Intermediate Value Theorem.

(h) Explain what are meant by a *path* and a *path-connected* space.

(i) Show that certain spaces are path-connected.

(j) State and make use of properties of path-connected spaces.

(k) Appreciate the links and differences between path-connected and connected spaces.

(l) Describe an example of a metric space that is connected, but not path-connected.

# Solutions to exercises

## Solution to Exercise 1.1

One homeomorphism from $A$ to $B$ is the function $f\colon A \to B$ given by $f\colon x \mapsto nx$ (which is clearly a $(d_A^{(1)}, d_B^{(1)})$-continuous bijection). The number $n$ does not have to be a natural number, but $n$ must not be 0 since we cannot have a bijection between a point and an interval.

## Solution to Exercise 1.2

One such map is the linear map that takes $a$ to $c$ and $b$ to $d$ and stretches (or shrinks) the first interval to cover the second. This is the function $f_I\colon X \to Y$ given by $f_I\colon x \mapsto c + \frac{d-c}{b-a}(x-a)$. As $f_I$ is a linear map, it is clearly a bijective $(d_X^{(1)}, d_Y^{(1)})$-continuous map with a $(d_Y^{(1)}, d_X^{(1)})$-continuous inverse.

It does not matter if $a < b$ and $c > d$ since the homeomorphism can 'flip the interval over'. This action is still continuous. For example, $f\colon x \mapsto -x$ gives a homeomorphism from $[0,1]$ to $[-1,0]$.

## Solution to Exercise 1.3

The map $f_I\colon x \mapsto c + \frac{d-c}{b-a}(x-a)$ used in Exercise 1.2, when restricted to the half-open interval, gives a homeomorphism from $X$ to $Y$. The intervals $(a,b]$ and $[c,d)$ are homeomorphic since the map $f\colon x \mapsto -x$ gives a homeomorphism from $[a,b)$ to $(-b,-a]$. So any half-open interval of the form $[a,b)$ can be 'flipped over' to one of the form $(x,y]$ before applying a homeomorphism of the form $f_I$ as described earlier.

## Solution to Exercise 1.4

## Solution to Exercise 1.5

The interval $(a,b]$ incudes $b$. Therefore we must have $f(b) \in (c,d)$. However, $f(b) = d$ and $d \notin (c,d)$ so $f$ is not a map from $(a,b]$ to $(c,d)$.

## Solution to Exercise 1.6

Since one set is infinite and the other is finite, there can be no bijection between them and so no homeomorphism.

## Solution to Exercise 1.7

Consider the function $\phi\colon \mathbb{R} \to (a,\infty)$ given by $\phi(x) = a + \exp(x)$. This is a sum of basic continuous functions (so is continuous) and has a continuous inverse (namely $\phi^{-1}\colon (a,\infty) \to \mathbb{R}$ given by $\phi^{-1}(x) = \log(x-a)$). The fact that the inverse is defined shows that $\phi$ is onto, and standard properties of $\exp(x)$ show that it is one–one.

## Solution to Exercise 2.1

There are many possible solutions. One such is $X = (0,1) \cup \{2\}$, with the metric induced by the Euclidean metric on $\mathbb{R}$, $d_X^{(1)}$. For then $(0,1) = [0,1] \cap \big((0,1) \cup \{2\}\big)$ and so, by Theorem 3.12 of Chapter 16, $(0,1)$ is $d_X^{(1)}$-closed in $X = (0,1) \cup \{2\}$.

## Solution to Exercise 2.2

Note that these solutions are not unique.

(a) We are working in the subspace $(X,d)$ where $X = (-\infty,0) \cup (1,\infty)$ and $d$ is the metric induced by the Euclidean metric on $\mathbb{R}$. Therefore we know that the $d$-closed sets in $X$ will be the intersections of $d^{(1)}$-closed sets in $\mathbb{R}$ with $X$. So to show that $(-\infty,0)$ is $d$-closed, we need to find a $d^{(1)}$-closed set $A$ in $\mathbb{R}$ such that $A \cap X = (-\infty,0)$. Now $(-\infty,\frac{1}{2}]$ is $d^{(1)}$-closed in $\mathbb{R}$ and $(-\infty,0) = (-\infty,\frac{1}{2}] \cap X$, and so is $d$-closed when considered as a subset of $X$. (We could have used any set $(-\infty,x]$ with $x \in [0,1)$.) Also $(-\infty,0) = (-\infty,\frac{1}{2}) \cap X$ and so is $d$-open (as $(-\infty,\frac{1}{2})$ is $d^{(1)}$-open in $\mathbb{R}$).

(b) Here we are working in the subspace $(X,d)$ where $X = (-\infty,0] \cup (1,\infty)$ and $d$ is the induced metric by the Euclidean metric on $\mathbb{R}$. Therefore we know that the $d$-closed sets in $X$ will be the intersections of $d^{(1)}$-closed sets in $\mathbb{R}$ with $X$. The interval $(-\infty,0] = (-\infty,\frac{1}{2}] \cap X$ and so is $d$-closed (as $(-\infty,\frac{1}{2}]$ is $d^{(1)}$-closed in $\mathbb{R}$). Similarly, $\{x \in \mathbb{R} \colon -\infty < x \leq 0\} = (-\infty,\frac{1}{2}) \cap X$ and so is $d$-open (as $(-\infty,\frac{1}{2})$ is $d^{(1)}$-open in $\mathbb{R}$).

(c) This time we are working in $(X,d)$ where $X = \{0\} \cup (3,4]$ and $d$ is the metric induced by the Euclidean metric on $\mathbb{R}$, and so we need to consider the intersections of $d^{(1)}$-closed and $d^{(1)}$-open subsets of $\mathbb{R}$ with $X$. The interval $[-1,1]$ is $d^{(1)}$-closed in $\mathbb{R}$ and the intersection of this interval with $X$ is $\{0\}$, and so $\{0\}$ is $d$-closed in $X$. Also $\{0\} = (-1,1) \cap X$ and so is $d$-open in $X$ (as $(-1,1)$ is $d^{(1)}$-open in $\mathbb{R}$).

(d) First we want to show that $(0,1)$ is $d$-closed as a subset of $\mathbb{R} - \mathbb{Z}$ where $d$ is the metric induced by the Euclidean metric on $\mathbb{R}$. As we are working with the induced metric, we need to consider the intersections of $d^{(1)}$-closed sets in $\mathbb{R}$ with $\mathbb{R} - \mathbb{Z}$. Now $[0,1]$ is $d^{(1)}$-closed in $\mathbb{R}$ and so $(0,1) = [0,1] \cap (\mathbb{R} - \mathbb{Z})$ is $d$-closed in $\mathbb{R} - \mathbb{Z}$. Also $(0,1) = (0,1) \cap (\mathbb{R} - \mathbb{Z})$ and so is $d$-open in $\mathbb{R} - \mathbb{Z}$.

## Solution to Exercise 3.1

(a) Here $X = (-\infty, 0) \cup (1, \infty)$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$. We need to find two non-empty disjoint sets, each of which is simultaneously $d_X$-closed and $d_X$-open as a subset of $(X, d_X)$, and whose union is the whole of $X$. We make the natural choices, $U = (-\infty, 0)$ and $V = (1, \infty)$, which we know are each simultaneously $d_X$-closed and $d_X$-open from Exercise 2.2. Therefore a disconnection is $\{(-\infty, 0), (1, \infty)\}$.

(b) Here $X = (-\infty, 0] \cup (1, \infty)$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$. We need to find two non-empty disjoint sets, each of which is simultaneously $d_X$-closed and $d_X$-open as a subset of $(X, d_X)$, and whose union is the whole of $(-\infty, 0] \cup (1, \infty)$. We make the natural choices, $U = (-\infty, 0]$ and $V = (1, \infty)$, which we know are each simultaneously $d_X$-closed and $d_X$-open from Exercise 2.2. Therefore a disconnection is $\{(-\infty, 0], (1, \infty)\}$.

(c) Here $X = \{0\} \cup (3, 4]$ and $d_X$ is the metric induced by the Euclidean metric on $\mathbb{R}$. We need to find two non-empty disjoint sets, each of which is simultaneously $d_X$-closed and $d_X$-open as a subset of $(X, d_X)$, and whose union is the whole of $\{0\} \cup (3, 4]$. We make the natural choices, $U = \{0\}$ and $V = (3, 4]$, which we know are each simultaneously $d_X$-closed and $d_X$-open from Exercise 2.2. Therefore a disconnection is $\{\{0\}, (3, 4]\}$.

## Solution to Exercise 3.2

Since the only subset of $\emptyset$ is $\emptyset$ itself, it follows that there is only one possible subset – namely, $\emptyset$ – and so no disconnection exists. So this space is connected.

## Solution to Exercise 3.3

The only subsets of $X$ are $\emptyset$, $\{x\}$ and $X$. Thus we cannot find two disjoint proper subsets and no disconnection of $X$ can exist. Therefore $X$ is connected.

## Solution to Exercise 3.4

We look for a proper subset which is simultaneously $d_0$-closed and $d_0$-open. Since $d_0$ is the discrete metric, we know from Chapter 16 that all subsets of $X$ are $d_0$-closed. In particular, each of $\{x\}$ and $\{x\}^c$ is $d_0$-closed. As $\{x\}^c$ is $d_0$-closed $\{x\} = (\{x\}^c)^c$, being the complement of a $d_0$-closed set, is $d_0$-open. As $X$ contains at least two points, we know that $\{x\} \neq X$. Thus, the set $\{x\}^c$ is non-empty. Therefore $\{x\}$ is a $d_0$-closed and $d_0$-open proper subset of $X$, and so $X$ is disconnected by Theorem 3.2.

## Solution to Exercise 3.5

If $X$ is disconnected then, by definition, $X$ can be expressed as the union of two non-empty $d$-closed sets $E$ and $F$. Therefore we need to show that if $X$ can be expressed as the disjoint union of two $d$-closed sets $X = E \cup F$ with $E \cap F = \emptyset$, then $X$ can be expressed as the disjoint union of two non-empty sets, each of which is simultaneously $d$-closed and $d$-open. Assume that $X = E \cup F$ with $E$ and $F$ disjoint, non-empty and $d$-closed. Then $E = X - F$ and so $E$ is the complement of a $d$-closed set and is $d$-open. Similarly, $F$ is $d$-open as it is the complement of $E$. Therefore $E$ and $F$ are each simultaneously $d$-closed and $d$-open and $X$ is disconnected.

## Solution to Exercise 3.6

We look for a way to divide $\mathbb{R} - \mathbb{Q}$ into two disjoint, $d_{\mathbb{R}-\mathbb{Q}}$-closed subsets of $\mathbb{R} - \mathbb{Q}$. As $\mathbb{R} - \mathbb{Q}$ has the metric induced by the Euclidean metric on $\mathbb{R}$, we know that a $d_{\mathbb{R}-\mathbb{Q}}$-closed subset of $\mathbb{R} - \mathbb{Q}$ is the intersection of a $d^{(1)}$-closed subset of $\mathbb{R}$ with $\mathbb{R} - \mathbb{Q}$.

Let $q$ be a rational number, such as 42. Then $E = (-\infty, 42] \cap (\mathbb{R} - \mathbb{Q})$ and $F = [42, \infty) \cap (\mathbb{R} - \mathbb{Q})$ are two disjoint $d_{\mathbb{R}-\mathbb{Q}}$-closed sets whose union is $\mathbb{R} - \mathbb{Q}$. Therefore by Proposition 3.4 $(\mathbb{R} - \mathbb{Q}, d_{\mathbb{R}-\mathbb{Q}})$ is disconnected.

## Solution to Exercise 3.7

There are many possible solutions to this question. One is as follows. Let $X = \{0, 1\}$ with the discrete metric, which is disconnected by Exercise 3.4. Let $Y = \{1\}$ with the discrete metric, which is connected by Exercise 3.3. Define $f\colon X \to Y$ by $f(x) = 1$ for $x \in X$. Then $f$ is $(d, e)$-continuous (by Proposition 1.10). Also any discrete space with more than one point is disconnected (by Exercise 3.4) and so $f$ is an onto $(d_0, d_0)$-continuous map from the disconnected space $(X, d_0)$ to the connected space $(Y, d_0)$. (In fact, any map from a discrete space with at least two points onto a connected space will suffice here.)

## Solution to Exercise 3.8

Suppose that $\chi_U$ is $(d, d_0)$-continuous. The set $\{1\}$ is both $d_0$-closed and $d_0$-open in $\{0, 1\}$. Therefore $\chi_U^{-1}(\{1\}) = U$ is $d$-closed and $d$-open in $X$ (by Book D, Chapter 16, Theorem 7.1 and Corollary 7.2).

## Solution to Exercise 3.9

Suppose that $U$ is both $d$-closed and $d$-open. We need to show that the preimage of any $d_0$-closed subset of $\big(\{0, 1\}\big), d_0)$ is $d$-closed. Since $d_0$ is the discrete metric, every subset of the set is $d_0$-closed; and since it is a finite set, we can simply list the possible subsets and deal with each in turn. The subsets are $\emptyset, \{0\}, \{1\}, \{0, 1\}$. Now $\chi_U^{-1}(\emptyset) = \emptyset$, which is $d$-closed in $X$; $\chi_U^{-1}(\{0\}) = U^c$, which is $d$-closed in $X$; $\chi_U^{-1}(\{1\}) = U$ and $\chi_U^{-1}(\{0, 1\}) = X$, which are both $d$-closed sets in $X$. Therefore $\chi_U$ is a $(d, d_0)$-continuous map by Book D, Chapter 16, Theorem 7.1.

## Solution to Exercise 3.10

There are (infinitely) many ways to answer this question. An example is: $\mathbb{Q} = \{q \in \mathbb{Q} : q < \sqrt{2}\} \cup \{q \in \mathbb{Q} : q > \sqrt{2}\}$ and $\mathbb{Q} = \{q \in \mathbb{Q} : q < \pi\} \cup \{q \in \mathbb{Q} : q > \pi\}$.

## Solution to Exercise 3.11

We write $d$ for the induced metric $d^{(1)}_{(\mathbb{R}-\mathbb{Z})}$. By Theorem 2.1 the interval $I_n = (n, n+1)$, where $n \in \mathbb{Z}$, is $d$-open in $X$ (as it is $d^{(1)}$-open in $\mathbb{R}$ and equal to the intersection of itself with $\mathbb{R}$). It is also $d$-closed (as it is equal to the intersection of the $d^{(1)}$-closed interval $[n, n+1]$ with $X$). Therefore $I_n = (n, n+1)$ is simultaneously $d$-open and $d$-closed, so cannot be a proper subset of a connected subset of $X$ (by Theorem 3.2). By the given assumption, $I_n$ is connected in $X$. Hence the intervals $I_n$ form the components of $X$.

## Solution to Exercise 3.12

When $(0, 1)$ has the discrete metric, each of the sets $\{x\}$, where $x \in (0, 1)$, is simultaneously $d_0$-open and $d_0$-closed (as every subset of any space with the discrete metric is simultaneously $d_0$-closed and $d_0$-open). Therefore any subset of $(0, 1)$ containing more than one element is disconnected (since if $x$ and $y$ are both in the same subset, $A$, then $\{x\}$ and $A - \{x\}$ would form a disconnection) and so the components must be the singleton sets $\{x\}$.

## Solution to Exercise 3.13

Let $(X, d)$ be a metric space, and let $x \in X$. Write $\mathrm{Cl}(X)$ for $\mathrm{Cl}_{(X,d)}(X)$.

It follows from Theorem 3.12 that $\mathrm{Cl}(C_d(x))$ is connected. Now $\mathrm{Cl}(C_d(x))$ contains $C_d(x)$. Also since $\mathrm{Cl}(C_d(x))$ is connected we have $\mathrm{Cl}(C_d(x)) \in \mathcal{F}_x = \{A \subseteq X : x \in A \text{ and } A \text{ is connected}\}$. Therefore $\mathrm{Cl}(C_d(x)) \subseteq \bigcup_{A \in \mathcal{F}_x} A$, which is equal to $C_d(x)$. So $\mathrm{Cl}(C_d(x)) = C_d(x)$, and hence $C_d(x)$ is closed.

## Solution to Exercise 4.1

Let $B = \{-a : a \in A\}$. Since $A$ is bounded below, $B$ is bounded above, and so $B$ has a least upper bound, $M$ say. Since $b \leq M$ for all $b \in B$, it follows that $-a \leq M$ for all $a \in A$, and so $-M \leq a$ for all $a \in A$. Hence $-M$ is a lower bound for $A$.

If $m' > -M$, then $-m' < M$, and so there is $b \in B$ for which $-m' < b$. But this means there is $a \in A$ with $-m' < -a$, and this rearranges to give $a < m'$. Hence $-M$ is the greatest lower bound of $A$.

## Solution to Exercise 4.2

By Worked Exercise 1.13, the interval $(0, 1)$ is homeomorphic to $\mathbb{R}$. By Corollary 3.7, connectedness is preserved by homeomorphism. Therefore $(0, 1)$ must be connected.

### Solution to Exercise 4.3

The interval $(0, 1)$ with the Euclidean metric is connected. Therefore there is a single component, namely $(0, 1)$ itself.

### Solution to Exercise 4.4

All these intervals contain $(a, b)$ and are contained in the interval $[a, b]$, the $d^{(1)}$-closure of $(a, b)$ in $\mathbb{R}$. Therefore they are connected, by Theorem 3.12.

### Solution to Exercise 4.5

A disconnection of $\mathbb{R} - \{a\}$ is $\{(-\infty, a), (a, \infty)\}$.

### Solution to Exercise 4.6

The continuous image of a connected set is connected by Lemma 3.6. The map $f : x \mapsto \big(\sin(x), \cos(x)\big)$ gives a $(d_I^{(1)}, d_C^{(2)})$-continuous map from the interval $I = [0, 2\pi)$ to $C$. Therefore $C$ is connected.

### Solution to Exercise 4.7

The proof is by contradiction. We know that $A$ and $B$ are both connected. Suppose that there is a homeomorphism $\phi\colon A \to B$. Let $a \in A$ be the point for which $\phi(a) = 0$ and $b \in A$ be the point for which $\phi(b) = 1$. The restriction rule for homeomorphisms implies that $A - \{a, b\}$ is homeomorphic to $\phi\big(A - \{a, b\}\big) = B - \{0, 1\} = (0, 1)$. This, however, is a contradiction since $A - \{a, b\}$ is disconnected, while $(0, 1)$ is connected. Thus $A$ is not homeomorphic to $B$.

Note that we need to consider both the point $a$ for which $\phi(a) = 0$ and the point $b$ for which $\phi(b) = 1$ to ensure a disconnection here. This is because it is possible that $\phi(1) = 0$ or $\phi(1) = 1$, in which case $A - \{1\} = (0, 1)$, which is connected. The second point ensures that we have a point interior to $(0, 1)$ with which to disconnect $A$.

### Solution to Exercise 4.8

Let $(X, d_X)$ and $(Y, d_Y)$ be non-empty metric spaces, and suppose that $X \times Y$ is $e$-connected. We know that the projection functions are continuous, and that $X = p_X(X \times Y)$, $Y = p_Y(X \times Y)$. So $X$ and $Y$ are images of a connected space under continuous mappings, and so are connected (by Lemma 3.6).

### Solution to Exercise 4.9

The function $f$ is continuous and the space $(\mathbb{R}^2, d^{(2)})$ is connected, so we can apply Theorem 4.6, the Intermediate Value Theorem. We must find a point in $\mathbb{R}^2$ at which the function takes a value less than 42 and another point at which the function takes a value greater than 42. For example,

$$f(0, 0) = 0 \quad \text{and} \quad f(2, 0) = 5 \times 2^5 = 160.$$

So, by the Intermediate Value Theorem, there is a point in the plane at which $f$ attains the value 42. (The theorem does not tell us what this point is, just that such a point exists, since the Intermediate Value Theorem is an existence result.)

See the discussion after Book D, Chapter 13, Worked Exercise 2.13.

## Solution to Exercise 5.1

By Theorem 7.1 of Chapter 16, the constant function $p\colon [0,1] \to X$ given by $p(t) = x$ will be $(d^{(1)}_{[0,1]}, d)$-continuous if, and only if, $p^{-1}(E)$ is $d^{(1)}_{[0,1]}$-closed whenever $E$ is a $d$-closed set in $X$. So let $E$ be a $d$-closed set in $X$. Then either $x \in E$ or $x \notin E$. If $x \in E$ then $p^{-1}(E) = [0,1]$, which is $d^{(1)}_{[0,1]}$-closed. If $x \notin E$ then $p^{-1}(E) = \emptyset$, which is also $d^{(1)}_{[0,1]}$-closed. Therefore we have the result.

## Solution to Exercise 5.2

Define $f\colon [0,1] \to X$ by $f(t) = (t, \frac{1}{2}t)$. We note that $(p_1 \circ f)(t) = t$ and $(p_2 \circ f)(t) = \frac{1}{2}t$ are $(d^{(1)}_{[0,1]}, d^{(1)}_{[0,1]})$-continuous functions from $[0,1]$ to $[0,1]$. So, by Theorem 3.12 of Chapter 15, the map $f$ is $(d^{(1)}_{[0,1]}, d^{(2)}_X)$-continuous. Hence, since $f(0) = (0,0)$ and $f(1) = (1, \frac{1}{2})$, the map $f$ gives a path joining $(0,0)$ and $(1, \frac{1}{2})$.

## Solution to Exercise 5.3

In this solution, all metrics are the metrics induced by the Euclidean metric on $\mathbb{R}$.

Define $q\colon [0,1] \to X$ by $q(t) = p(1-t)$. Then $q(0) = p(1) = b$ and $q(1) = p(0) = a$. Moreover, since $q = p \circ f$, where $f\colon [0,1] \to [0,1]$ is the basic continuous function given by $f(t) = 1 - t$, it follows that $q$ is a composition of continuous functions and so is continuous. Hence $q$ is a path in $X$ with initial point $b$ and final point $a$.

## Solution to Exercise 5.4

If $c \neq d$ are two points in $[a,b]$ then we need to show that there is a path within $[a,b]$ from $c$ to $d$.

If $a = b$, we are done. Assume (without loss of generality) that $a < b$ and let $c, d \in [a,b]$. If $c = d$, we are done, so we assume (without loss of generality) that $c < d$. We define $P\colon [0,1] \to [a,b]$ by

$$P(t) = (1-t)c + td.$$

Note that
$(1-t)c + td \leq (1-t)b + tb = b$
and
$(1-t)c + td \geq (1-t)a + ta = a$.

We know from Worked Exercise 5.6 that $P$ defines a path from $c$ to $d$ in $\mathbb{R}$. Thus, using Theorem 1.6 in Chapter 15, we can deduce that $P$ defines a path in $[a,b]$ if $P([0,1]) \subseteq [a,b]$. This is clearly true 'by inspection'. Therefore $[a,b]$ is path-connected.

## Solution to Exercise 5.5

The metrics have been added in to the relevant section below.

*Continuity:* let $p_j\colon \mathbb{R}^n \to \mathbb{R}$ be the $j$th projection function, defined by $p_j(\mathbf{a}) = a_j$. Then $(p_j \circ P)(t) = (1-t)a_j + tb_j$, so $p_j \circ P$ is linear and so is clearly $(d^{(1)}_{[0,1]}, d^{(1)})$-continuous. Since $p_j \circ P$ is $(d^{(1)}_{[0,1]}, d^{(1)})$-continuous for $j = 1, 2, \ldots, n$, it follows from the generalisation of Theorem 3.12 of Chapter 15 that $P$ is $(d^{(1)}_{[0,1]}, d^{(n)})$-continuous.

## Solution to Exercise 5.6

For any $x \in [0,1]$, and any $s, t \in [0,1]$,

$$\begin{aligned} |P(t)(x) - P(s)(x)| &= |(1-t)f(x) + tg(x) - (1-s)f(x) - sg(x)| \\ &= |(t-s)(g(x) - f(x))| \\ &= |t-s|\,|g(x) - f(x)| \\ &\leq |t-s| d_{\max}(f,g). \end{aligned}$$

Taking the maximum over $x \in [0,1]$, we have

$$d_{\max}(P(t), P(s)) \leq |t-s| d_{\max}(f,g).$$

Therefore $P$ is a $(d^{(1)}_{[0,1]}, d_{\max})$ Lipschitz function and so is $(d^{(1)}_{[0,1]}, d_{\max})$-continuous (by Theorem 3.2 of Chapter 15).

## Solution to Exercise 5.7

Suppose that there is a homeomorphism $\phi\colon \mathbb{R} \to \mathbb{R}^2$. Then removing a point, $p$, from $\mathbb{R}$ and removing $\phi(p)$ from $\mathbb{R}^2$ would give us a homeomorphism from $\mathbb{R} - \{p\}$ to $\mathbb{R}^2 - \{\phi(p)\}$. This is impossible since $\mathbb{R} - \{p\}$ is not connected (and so is not path-connected), but $\mathbb{R}^2 - \{p\}$ is path-connected (by Worked Exercise 5.15) and path-connectedness is a topological invariant by Corollary 5.9.

## Solution to Exercise 5.8

We know, from Exercise 3.4, that $X$ is not connected. It follows from Theorem 5.13 that $X$ cannot be path-connected.

## Solution to Exercise 6.1

For each $n$, let $l_{n,1}$ be the map $l$ restricted to the set: $I_{n,1} = \left\{x \in \mathbb{R}\colon \frac{1}{n+1} \leq x \leq \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right)\right\}$ and $l_{n,2}$ be the map $l$ restricted to the set $I_{n,1} = \left\{x \in \mathbb{R}\colon \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right) \leq x \leq \frac{1}{n},\ n \in \mathbb{N}\right\}$. Both maps $l_{n,i}$ are linear and so are $(d^{(1)}_{I_{n,i}}, d^{(1)})$-continuous. Also the maps $l_{n,i}$ agree on appropriate points. Therefore $l$ is continuous by Lemma 5.3.

Let $p_1$ and $p_2$ be the projection functions from $\mathbb{R}^2$ to $\mathbb{R}$. Then $(p_1 \circ L)(x) = x$ and $(p_2 \circ L)(x) = l(x)$, and so $p_1 \circ L$ and $p_2 \circ L$ are both $(d^{(1)}_{[0,1]}, d^{(1)})$-continuous functions. Hence $L$ is $(d^{(1)}_{[0,1]}, d^{(2)})$-continuous, by Proposition 1.19 of Chapter 14.

## Solution to Exercise 6.2

(a) Since $0 \leq b \leq 1$, we have:

$$\frac{1}{n+1} \leq x_n \leq \frac{2n+1}{2n(n+1)} = \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right),$$

and so $l(x_n) = 2n(n+1)x_n - 2n = (b+2n) - 2n = b$.

(b) We want $\mathbf{a}_n$ to be inside $B_{d_C}(\mathbf{b}, r)$. This occurs if the $x$-coordinate, $x_n$ is less than $r$.

$$x_n = \frac{b+2n}{2n(n+1)}$$

so we need to find $N$ with

$$\frac{b+2N}{2N(N+1)} < r.$$

Choosing any $N$ such that $2N((N+1)r-1)) > b$, which always exists, will suffice.

(c) Part (a) shows that we can find a sequence of points $(\mathbf{a}_n)$, where $\mathbf{a}_n \in A_n$ and where the $x$-coordinate of each point is $b$. Part (b) shows that for any $r > 0$ we can find an $N \in \mathbb{N}$ such that $\mathbf{a}_n \in B_{d_C}(\mathbf{b}, r)$. In other words, we have shown that $\mathbf{a}_n$ is $d_C$-convergent to $\mathbf{b}$. Therefore $\mathbf{b}$ is a limit point of $(\mathbf{a}_n)$ and $\mathbf{b} \in \mathrm{Cl}_{(C,d_C)}(A)$. Thus $\mathrm{Cl}_{(C,d_C)}(A) = C$.